理解科学丛书·卢昌海科普著作

FROM
SINGULARITIES
TO WORMHOLES

从奇点到虫洞

广义相对论专题选讲

卢昌海◎著

奇点、黑洞、白洞、虫洞、时间旅行……

对这些科普和科幻作品中迷人概念的深度探索；

霍金、彭罗斯、威顿、丘成桐、索恩……

对这些著名科学家精彩工作的细致解读。



清华大学出版社

目　录

作者简介

卢昌海，出生于杭州，本科就读于复旦大学物理系，毕业后赴美留学，于2000年获美国哥伦比亚大学物理学博士学位，目前旅居纽约。著有《那颗星星不在星图上：寻找太阳系的疆界》、《太阳的故事》和《黎曼猜想漫谈》，并曾在《中国青年报》、《数学文化》、《科幻世界》、《现代物理知识》、《中学生天地》、《科学画报》等报纸、杂志上发表几十篇科普及专业科普作品。

FROM

SINGULARITIES

TO WORMHOLES

从奇点到虫洞
——广义相对论专题选讲

卢昌海　著

清华大学出版社
北　京

内容简介

本书以能量条件为线索，介绍普通教材中较少涉及或较少深入介绍的若干广义相对论专题，包括奇点及奇点定理、正质量定理、宇宙监督假设与虫洞物理学等。本书的介绍既包含背景或历史的介绍，也不回避必要的数学细节，融语言的生动风趣与内容的严谨翔实于一体，可使读者获得对这些专题的较深入的了解。

本书的介绍可作为普通教材的补充，供大学生、研究生及其他具有一定基础的物理爱好者阅读。

图书在版编目（CIP）数据

从奇点到虫洞：广义相对论专题选讲 / 卢昌海著．—北京：清华大学出版社，2013

（理解科学丛书）

ISBN 978-7-302-32739-4

Ⅰ．①从…　Ⅱ．①卢…　Ⅲ．①广义相对论　Ⅳ．①O412.1

中国版本图书馆CIP数据核字（2013）第130823号

责任编辑：邹开颜　赵从棉

封面设计：蔡小波

责任校对：赵丽敏

责任印制：杨　艳

出版发行：清华大学出版社

网　址：http://www.tup.com.cn，http://www.wqbook.com

地　址：北京清华大学学研大厦A座

邮　编：100084

社总机：010-62770175

邮　购：010-62786544

投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn

质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn

印装者：三河市金元印装有限公司

经　销：全国新华书店

开　本：165mm×240mm

印　张：9.75

字　数：136千字

版　次：2013年12月第1版

印　次：2013年12月第1次印刷

产品编号：048341-01

第1章 能量条件

1.1 引言

大家知道，广义相对论的场方程（即爱因斯坦场方程）[1]

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi T_{\mu\nu} \tag{1.1.1}$$

是一组有关时空度规的二阶非线性偏微分方程，求解这样的方程组是极其困难的。在20世纪60年代初以前，物理学家们对爱因斯坦场方程的很大一类研究都局限于在各种各样的简化条件——比如特定的对称性——下求解场方程。在这方面最著名的成果之一是德国物理学家施瓦西（Karl Schwarzschild，1873—1916）于1916年得到的施瓦西解，其度规为

$$ds^2 = \left(1 - \frac{2m}{r}\right)dt^2 - \left(1 - \frac{2m}{r}\right)^{-1}dr^2 - r^2 d\Omega^2 \tag{1.1.2}$$

其中m为质量参数。另一个同样著名的成果则是俄国物理学家弗里德曼（Alexander Friedmann，1888—1925）于1922年得到的弗里德曼解，其度规被称为罗伯逊-沃尔克度规（Robertson-Walker metric），为[2]

$$ds^2 = dt^2 - R^2(t)\left(\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2\right) \tag{1.1.3}$$

其中R为标度因子，k取值为0、-1或1，分别对应于平直（flat）、负常曲率（constant negative curvature）及正常曲率（constant positive curvature）空间。这两个度规分别是广义相对论在天体物理学和宇宙学上应用最为广泛的度规。

但这两个解的发现也带来了一个共同的问题，那就是它们所对应的

度规均具有奇异性。施瓦西度规是一个静态度规，它的奇异性（由上述表达式可以很容易地看到）出现在r＝0及r＝2m处。这其中r＝2m处的奇异性（一度被称为施瓦西奇点）后来被证明只是坐标选择导致的表观奇异性，可以通过坐标变换予以消除〔3〕；而r＝0处的奇异性则是真正的物理奇点，时空曲率在趋近该点时趋于发散〔4〕。这个奇点被称为曲率奇点。罗伯逊-沃尔克度规由于是动态度规，情形稍微复杂些。当k＝1（即空间具有正曲率）时这一度规在r＝1处似乎具有奇异性，但这也是坐标选择导致的表观奇异性（读者们不妨自己寻找一个坐标变换来消去这一表观奇异性）。除去这一表观奇异性，从形式上看罗伯逊-沃尔克度规似乎没有其他显而易见的奇异性。但把这一度规代入到场方程中，研究它的动力学演化就会发现，对于我们观测到的膨胀宇宙来说，只要宇宙当前的物质分布满足一个很宽泛的条件，罗伯逊-沃尔克度规中的标度因子R（t）在过去某个有限时刻就必定等于零。在那个时刻（通常定义为t＝0）宇宙的空间线度为零，物质密度则发散，因此那也是一个真正的物理奇点，被称为宇宙学奇点，或大爆炸（The Big Bang）。

这些奇点的出现是物理学家们所不乐意见到的，因为物理世界中并不存在真正意义上的无穷大。对于一个物理理论来说，出现无穷大往往意味着它的失效。因此奇点的出现对广义相对论是一种危机。不过当时物理学家们所知道的爱因斯坦场方程的解的数量十分有限，而且那些解大都具有很高的对称性（因为只有那种情形下的场方程才容易求解），比如施瓦西解具有球对称性，弗里德曼解则是均匀及各向同性的。这就给物理学家们提出了一个问题：由这几个特殊解所展示的危机究竟有多大的普遍性？或者说奇点的出现会不会只是那几个特殊解所具有的特殊对称性导致的特殊效应？如果是的话，那情势就不算太严重，因为那些对称性在现实世界里是不可能绝对严格地实现的，从而危机也就不具有普遍性。20世纪60年代，物理学家们对这一问题有两种不同的看法：一

种看法认为奇点的出现确实只是特殊对称性导致的特殊效应，如果考虑一般（即没有严格对称性）的情形，奇点将不会出现。持这种观点的代表人物是苏联物理学家栗弗席兹（Evgeny Lifshitz，1915—1985）、卡拉特尼科夫（Isaak Markovich Khalatnikov，1919— ）和贝林斯基（Vladimir Belinski，1941— ）等。与之相反的另一种看法则认为奇点在广义相对论中的出现是有普遍性的，从而并不是特殊对称性导致的特殊效应。持这种观点的代表人物是英国物理学家彭罗斯（Roger Penrose，1931— ）和霍金（Stephen Hawking，1942— ）等[5]。

这两组物理学家在奇点问题上不仅观点迥异，而且在这一领域所采用的研究方法也很不相同。栗弗席兹等人由于相信奇点的出现跟具体的解——尤其是其中的对称性——有关，因此把主要精力放在了求解一般——即没有严格对称性——情形下的场方程，以便探讨并检验在那种情形下是否不存在奇点；而彭罗斯和霍金等人也许是由于不认为奇点的出现跟具体的解有关，因此并不着眼于求解场方程，而大量运用了微分几何手段，通过所谓的“全局方法”（global method），在不直接求解场方程的情况下对奇点及奇点产生的条件进行了系统分析。如果说栗弗席兹等人的方法是正面强攻，那么彭罗斯和霍金等人的方法则属于旁敲侧击。经过几年的努力，两种方法分出了高下。栗弗席兹等人的正面强攻收效不大——因为爱因斯坦场方程实在太复杂了。虽然栗弗席兹等人的胃口并不贪婪，他们只研究宇宙学奇点t＝0附近的解而非全局性解，同时也并不奢望精确求解而采用了近似手段，但在不具有对称性的情形下，他们的努力依然遭到了难以逾越的困难[6]。另一方面，彭罗斯和霍金等人的“旁敲侧击”却获得了极大的成功，他们证明了一系列被称为奇点定理（singularity theorem）的著名结果，成为了经典广义相对论中登峰造极的成果之一。

不过彭罗斯和霍金等人的方法虽然不需要直接求解场方程，却也并非“不食人间烟火”，因为它与物质能量动量张量的性质依然有着密切关

系。这一点从物理上讲是显而易见的，因为正是物质能量动量张量的分布决定了时空的结构。爱因斯坦曾经把他的场方程比喻为一座建筑，这座建筑的一半是用精美的大理石（fine marble）砌成的，另一半却是用劣质的木料（low-grade wood）建造的。用精美的大理石砌成的那一半是方程的左端：$R_{\mu\nu}-\frac{1}{2}g_{\mu\nu}R$，那是一个描述时空结构的优美的几何量，被称为爱因斯坦张量。而用劣质的木料建造的那一半则是方程的右端，也就是描述物质分布的能量动量张量：$8\pi T_{\mu\nu}$。为什么说这部分是用劣质木料建造的呢？因为自然界的物质分布种类繁多，物态方程千差万别，找不到一个普适的能量动量张量来描述所有已知的物质分布。不仅如此，在广义相对论所涉及的许多极端条件——比如某些星体内部的超高温、超高压、超高密度，宇宙演化的早期，以及引力坍缩的后期等条件——下还可能存在大量迄今未知的物质形态及分布。而且所有这些物质分布还可能在空间及时间上相互混合及转变。由于存在如此高度的复杂性，与爱因斯坦张量所具有的完全确定的数学结构相比，我们有关能量动量张量的知识无疑是极其贫乏的。

那么，在这种贫乏的知识下，如何才能研究诸如奇点的产生条件那样与物质的形态及分布密切相关，同时又具有很大普遍性的课题呢？彭罗斯和霍金等人采用了一种很高明的手法，那就是虽然谁也无法写下一个具有普适性的能量动量张量，但这一张量应当具备的某些基本条件（比如说能量密度必须大于等于零）在当时看来是具有很大的普适性的，因此他们假定物质的能量动量张量满足那些基本条件。另一方面，他们所使用的全局方法的威力之一就在于，只要利用那些基本条件，无须知道能量动量张量的具体形式，就可以得到许多非常有价值的结果。那些结果便是他们所证明的一系列奇点定理。而那些附加在能量动量张量上的条件则被统称为能量条件（energy condition）。如果说能量动量张量是用劣质木料建造的，那么能量条件的引进就好比是对那些劣质木

料套上几道铁箍进行加固，使它比原先的松散形式来得结实耐用。

1.2 能量条件

在本节中，将对几种主要的能量条件作一个简单介绍。不过，在介绍能量条件之前，我们首先要对能量动量张量本身的形式作以简单分析。在广义相对论研究中，为了让度规张量的形式尽可能简化，人们常常引进所谓的正交标架场（tetrad），对能量动量张量的分析也是如此。正交标架场（以下简称标架场）由一组正交归一的基矢量场$(e_a)^\mu$组成，其中拉丁字母a，b，…标识标架场的基矢量，希腊字母μ，ν，…表示基矢量的时空指标[7]。标架场的基矢量满足下列正交归一条件：

$$\eta^{ab}(e_a)^\mu(e_b)^\nu = g^{\mu\nu}, \quad g_{\mu\nu}(e_a)^\mu(e_b)^\nu = \eta_{ab} \qquad (1.2.1)$$

很明显，标架场不是唯一的，对一个标架场作局域洛伦兹变换所得到的仍然是标架场。由于洛伦兹群具有旋量表示[8]，因此标架场在讨论引力场与旋量场的相互作用时是非常重要的工具。对于我们所要讨论的能量条件来说，标架场的优点在于能量动量张量在标架场中的分量具有明确的测量意义。

霍金曾经把标架场下的能量动量张量分为四种类型，每种类型均可通过标架场中的洛伦兹变换约化为一个正则形式（canonical form）。这其中最重要的是第Ⅰ类，其正则形式为

$$T^{ab} = \mathrm{diag}(\rho, p_1, p_2, p_3) \qquad (1.2.2)$$

其中diag表示对角矩阵，ρ为标架场中的静止观测者（即世界线的切线沿基矢e_0方向的观测者）测量到的能量密度，p_i（i＝1，2，3）则为沿三个正交空间方向的主压强。除极少数特殊情形外，这种类型的能量动量张量涵盖了几乎所有物理上有意义的物质分布，我们在本书中将只讨论这种类型。

熟悉线性代数的读者可能会提出这样一个问题：第Ⅰ类能量动量张

量的正则形式其实就是该张量的对角化，但能量动量张量是一个实对称张量，按照线性代数中熟知的定理，实对称张量必定可以通过正交变换对角化，既然如此，能量动量张量岂不都应该是第Ⅰ类的？为什么在霍金的分类中它只是四种类型之一呢？这其中的原因在于普通线性代数所讨论的内积空间具有正定度规（positive definite metric），而广义相对论中的时空度规不是正定的（请读者想一想，度规的非正定性是如何破坏线性代数中有关实对称张量对角化的定理的？）。

下面我们就对几种主要的能量条件进行简单介绍：

弱能量条件（weak energy condition）：对所有类时矢量V_a，$T^{ab}V_aV_b \geq 0$。

利用T^{ab}的正则形式，我们可以证明：弱能量条件等价于**$\rho \geq 0$**及**$\rho + p_i \geq 0$（i＝1，2，3）**。充分性的证明非常简单：取$V_a = e_0$（即静止观测者）可得$\rho \geq 0$；取$V_a \to e_0 + e_i$（注意V_a是趋于而非等于$e_0 + e_i$，因为后者是类光的）则可得$\rho + p_i \geq 0$。接下来再证明必要性：假设$\rho \geq 0$及$\rho + p_i \geq 0$，则

$$T^{ab}V_aV_b = \rho V_0^2 + \sum_i p_i V_i^2 \geq \rho\left(V_0^2 - \sum_i V_i^2\right) \geq 0 \qquad (1.2.3)$$

其中第一个“≥”用到了$\rho + p_i \geq 0$，第二个“≥”用到了$\rho \geq 0$及V_a类时。

在弱能量条件中最重要的部分是$\rho \geq 0$，它表明能量密度处处为正。需要提醒读者注意的是，虽然上述推导是在使正则形式成立的特殊标架场中进行的，但$\rho \geq 0$这一结果适用于沿任意类时世界线运动的观测者所测得的能量密度（请读者想一想，这是为什么？）。由于物理上可以实现的所有观测者都是沿类时世界线运动的，因此弱能量条件表明任何物理观测者测得的能量密度都处处为正。

在弱能量条件中让V_a趋于类光，由能量条件的连续性可以得到：

零能量条件（null energy condition）：对所有类光矢量k_a，$T^{ab}k_a k_b \geq 0$。

显然（请读者自行证明），零能量条件等价于**$\rho + p_i \geq 0$（i＝1，2，3）**。零能量条件是一个非常弱的能量条件，比弱能量条件更弱。

强能量条件（strong energy condition）：对所有类时矢量V_a，$(T^{ab} - \frac{1}{2}g^{ab}T)V_a V_b \geq 0$。

由于爱因斯坦场方程可以改写为$R^{ab} = 8\pi G[T^{ab} - (1/2)g^{ab}T]$（其中$T = T^a_a$为能量动量张量的迹），因此强能量条件等价于一个几何条件$R^{ab}V_a V_b \geq 0$[9]。从物理上讲，强能量条件等价于**$\rho + \sum_i p_i \geq 0$**及**$\rho + p_i \geq 0$（i＝1，2，3）**。这一点的证明非常简单，只需注意到在正则形式下：

$$T^{ab} - \frac{1}{2}g^{ab}T = \frac{1}{2}\begin{pmatrix} \rho + \sum_i p_i & & & \\ & \rho + 2p_1 - \sum_i p_i & & \\ & & \rho + 2p_2 - \sum_i p_i & \\ & & & \rho + 2p_3 - \sum_i p_i \end{pmatrix} \tag{1.2.4}$$

然后效仿前面有关弱能量条件的证明即可（请读者自行推导上式并完成论证）。

显然，强能量条件比零能量条件强。但是与强、弱二字的正常含义不符的是，强能量条件与弱能量条件互不包含，而非前者强于后者。事

实上，多数物质的主压强p_i是正的，对于那些物质，强能量条件其实比弱能量条件还弱[10]。

主能量条件（dominant energy condition）：对所有类时矢量V_a，$T^{ab}V_aV_b \geq 0$，并且$T^{ab}V_b$非类空。

这个能量条件是在弱能量条件之上增添了能流密度矢量$T^{ab}V_b$非类空这一额外限制。在正则形式下这一额外限制可以表述为：$\|T^{ab}V_b\| = \rho^2 V_0^2 - \sum_i p_i^2 V_i^2 \geq 0$。取$V_b \to e_0 + e_i$可得$\rho^2 \geq p_i^2$。这比弱能量条件中的$\rho + p_i \geq 0$要强。为了证明$\rho^2 \geq p_i^2$也是保证额外限制成立的充分条件，只需注意到

$$\| T^{ab} V_b \| = \rho^2 V_0^2 - \sum_i p_i^2 V_i^2 \geqslant \rho^2 \left(V_0^2 - \sum_i V_i^2 \right) \geqslant 0 \qquad (1.2.5)$$

这里第一个“≥”用到了$\rho^2 \geq p_i^2$，第二个“≥”用到了$\rho \geq 0$及V_b类时。将这一结果附加到弱能量条件上可得：主能量条件等价于$\boldsymbol{\rho \geq |p_i|}$（**i＝1，2，3**）。从定义及上述结果中均可看出，主能量条件比弱能量条件强（从而也比零能量条件强）。但它与强能量条件互不包含。

看到这里，有些读者可能会产生这样一个疑问：主能量条件中的额外限制是说能流密度矢量非类空。我们知道，在相对论中如果一个四维矢量类空，就必定可以找到一个参照系，使该矢量的时间分量为负。对于能流密度矢量来说，时间分量就是能量密度，因此如果能流密度矢量类空，就说明必定存在一个参照系，在其中能量密度为负。但弱能量条件已经表明任何物理观测者测得的能量密度都处处为正，这岂不等于排除了能流密度矢量类空的可能性？如果这样的话，主能量条件中的额外限制变成了弱能量条件的推论，这两种能量条件岂不就变成等价的了？这种推理显然是错误的，但它究竟错在哪里呢？有兴趣的读者不妨思考

一下，以加深对能量条件及其观测意义的理解。

迹能量条件（trace energy condition）：T≡T^{a}_{a}≥0。

这是我们所要介绍的最后一种能量条件。它的表述与度规张量的符号约定有关，在本书中我们所用的约定是η_{ab}＝diag（1，-1，-1，-1）。如果做相反的约定，则迹能量条件的表述需改为T≤0。在正则形式下，迹能量条件等价于$\rho-\sum_{i}p_i\geq 0$，它与其他能量条件互不包含。

上面介绍的这几种能量条件有一个共同特点，那就是它们给出的都是每个时空点上能量动量张量所满足的条件，这样的能量条件被称为逐点能量条件（pointwise energy condition）。除逐点能量条件外，人们还常常使用另外一类能量条件，称为平均能量条件（average energy condition），它们给出的是能量动量张量沿特定的时空曲线（通常是类时或类光曲线）所满足的平均意义上的条件。平均能量条件比相应的逐点能量条件弱，因为它们允许逐点能量条件在局部意义上被破坏，只要这种破坏能被所涉及的时空曲线上其他区段的贡献所弥补即可。

在本书接下来的各专题中，我们将以能量条件为线索，介绍广义相对论中一些重要、优美或有趣的课题，比如奇点定理、正质量定理（positive mass theorem）、宇宙监督假设（cosmic censorship hypothesis）、虫洞物理学（wormhole physics）等。我们将会看到，能量条件在所有那些课题中都有着重要应用〔11〕。

注释

〔1〕在本系列的多数章节中，我们都取c＝G＝1的单位制，并且不考虑宇宙学项。

〔2〕在弗里德曼之后，比利时天文学家勒梅特（Georges Lemaître，1894—1966）、美国物理学家罗伯逊（Howard Robertson，1903—1961）、英国数学家沃尔克（Arthur Walker，1909

—2001）均对这个解做过研究，因此这个解所对应的度规在文献中有若干不同的名称，最常见的是罗伯逊-沃尔克（RW）度规，也有称为弗里德曼-罗伯逊-沃尔克（FRW）度规或弗里德曼-勒梅特-罗伯逊-沃尔克（FLRW）度规的。

〔3〕这一点最早是由勒梅特于1933年指出的。在那之前的1924年，英国天文学家爱丁顿（Arthur Eddington，1882—1944）曾找到过一个在r＝2m处非奇异的坐标，但他并未意识到自己结果的重要性。消除了所有表观奇异性的施瓦西度规的最大解析延拓则是由美国数学家克鲁斯卡（Martin Kruskal，1925—2006）和匈牙利数学家塞凯赖什（George Szekeres，1911—2005）于1960年各自独立地给出的。

〔4〕确切地讲，是由曲率张量构造出的某些标量——比如克莱茨曼标量（Kretschmann scalar）$R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}$——发散。这是与坐标选择无关的发散。

〔5〕对这两种观点的区分（以及将彭罗斯和霍金视为后一种观点的代表人物）散见于文献中。但彭罗斯和霍金是从一开始（即在得到理论证据之前）就认为奇点的出现具有普遍性，还是在彭罗斯证明了第一个奇点定理之后才这样认为？笔者猜测是前者，但在文献中未见到说明。

〔6〕栗弗席兹等人一度以为自己已对一般情况下奇点的不存在性作出了论述，并将结论写入了朗道（Lev Landau，1908—1968）与栗弗席兹的名著《经典场论》（The Classical Theory of Fields）。20世纪60年代末，随着奇点定理的出现，他们意识到了自己的错误，并托到访苏联的美国物理学家索恩（Kip Thorne，1940— ）以最快的速度将更正错误的文章秘密带到了西方。不过，栗弗席兹等人虽未能达到他们希望达到的目标，但他们的工作为后人研究奇点附近的时空性质提供了重要参考，依然功不可没。

〔7〕标架基矢（e_a）$^\mu$是时空坐标的函数，因此叫做标架场。“tetrad”这个名称通常是指四维的标架场（tetra这个词头的含意是“四”）。标架场的另一个常见的名称是vierbein，源于表示“四”的德语词头vier。在其他维数下，标架场还有其他一些常用名称，比如triad、pentad、funfbein、elfbein、vielbein等。

〔8〕切空间中的一般线性变换群GL（4，R）则没有旋量表示。

〔9〕这里不考虑宇宙学项。另外，其他能量条件也可以用类似方式改写成几何条件。

〔10〕有读者可能会问，既然强能量条件并不比弱能量条件强，那为什么会有这样的命名

呢？这是由于强能量条件可以写成$T^{ab}V_aV_b \geq$（1/2）T，而弱能量条件为$T^{ab}V_aV_b \geq 0$，由于通常T≥0，因此如果把这两个能量条件视为是对$T^{ab}V_aV_b$的约束条件，则强能量条件比弱能量条件强。当然，这种命名理由也是不严格的，因为T≥0本身就是一种能量条件（即迹能量条件），而非无条件成立的物理事实。

〔11〕事实上，从历史的角度讲，能量条件的提出本身正是由那些课题促成的，因为正是在研究那些课题的过程中，物理学家们意识到需要对能量动量张量附加一定的约束，这种约束就是能量条件。

第2章
奇点与奇点定理

2.1　什么是奇点？

自本节开始，我们将介绍能量条件在广义相对论中的若干应用，首先要介绍的是奇点定理。在广义相对论中，对奇点的研究是一个重要的课题，它既是能量条件最早的应用之一，也是全局方法在广义相对论中初试锋芒的范例。我们在1.1节中曾经提到，广义相对论的经典解——比如施瓦西解——存在奇异性。这其中有的奇异性——比如施瓦西解中的r＝2m——可以通过坐标变换予以消除，从而不代表物理上的奇点；而有的奇异性——比如施瓦西解中的r＝0——则是真正的物理奇点。很明显，在奇点研究中，真正的物理奇点才是我们感兴趣的对象。

那么究竟什么是广义相对论中真正的物理奇点（简称奇点）呢？

初看起来，这似乎是一个很简单的问题。奇点显然就是那些时空结构具有某种病态性质（pathological behavior）的时空点。但稍加推敲，就会发现这种说法存在许多问题。首先，“病态性质”是一个很含糊的概念，究竟什么样的性质是病态性质呢？显然需要予以精确化。其次，广义相对论与其他物理理论有一个很大的差异，那就是其他物理理论都预先假定了一个背景时空的存在，因此，那些理论如果出现奇点——比如电磁理论中点电荷所在处的场强奇点，我们可以明确地标识出奇点在背景时空中的位置[1]。但广义相对论所描述的是时空本身的性质。因此在广义相对论中一旦出现奇点，往往意味着时空本身的性质无法定义。另一方面，物理时空被定义为带洛伦兹度规（Lorentzian metric）的四维流形[2]，它在每一点上都具有良好的性质。因此，奇点的存在本身就是与物理时空的定义相冲突的，物理时空按定义就是没有奇点的，或者

换句话说，奇点并不存在于物理时空中[3]。

既然奇点并不存在于物理时空中，自然就谈不上哪一个时空点是奇点，从而也无法把奇点定义为时空结构具有病态性质的时空点了。但即便如此，像施瓦西解具有奇异性这样显而易见的事实仍然是无可否认的，因此关键还在于寻找一个合适的奇点定义。

为了寻找这样的定义，我们不妨想一想，为什么即便将施瓦西解中的r＝0那样的“麻烦制造者”排除在物理时空之外，我们仍然认为施瓦西解具有奇异性是显而易见的事实？答案很简单（否则就不叫显而易见了）：当一个试验粒子在施瓦西时空中沿径向落往中心（即r趋于0）时，它所接触到的时空曲率趋于发散。由于试验粒子的下落是沿非类空测地线进行的[4]，这启示我们这样来定义奇点：如果时空结构沿非类空测地线出现病态性质，则表明存在奇点。这一定义不涉及奇点的位置，从而不需要将奇点视为物理时空的一部分，也就避免了上面提到的与物理时空的定义之间的冲突。但是，这一定义还面临两个问题：一是“病态性质”这个含糊概念仍未得到澄清；二是在这一定义中，假如试验粒子沿非类空测地线需要经过无穷长的时间才会接触到时空结构的病态性质的话，奇点的存在就不具有观测意义。为了解决这两个问题，物理学家们提出了一个进一步的要求，即要求定义中涉及的非类空测地线具有有限“长度”，并且是不可延拓的（inextendible）[5]。这种具有有限的“长度”的不可延拓非类空测地线被称为不完备非类空测地线（incomplete non-spacelike geodesics）。

有了不完备非类空测地线这一概念，我们可以这样来定义奇点：如果存在不完备非类空测地线，则时空流形具有奇点。这就是多数广义相对论文献所采用的奇点定义。这种存在不完备非类空测地线的时空被称为非类空测地不完备时空，简称测地不完备时空（geodesically incomplete spacetime）。在一些文献中，按照不完备测地线的类型，还将测地不完备时空进一步细分为类时测地不完备时空与类光测地不完备

时空[6]。这一奇点定义的合理性体现在：在一个测地不完备的时空流形中，试验粒子可以沿不完备的非类空测地线运动，并在有限时间内从时空流形中消失。这种试验粒子在有限时间内从时空流形中消失的行为——即测地不完备性——可以视为是对时空结构具有“病态性质”这一含糊要求的精确表述。这样我们就既解决了“病态性质”的精确化问题，又使奇点具有了观测意义（即试验粒子在有限时间内就可以遇到奇点）。在一些文献中，还对奇点存在于过去还是未来进行区分：如果所涉及的非类空测地线是未来（过去）不可延拓的，则相应的奇点被称为未来（过去）奇点。

对于上述定义，还有一点需要补充。细心的读者可能注意到了，我们在引进“非类空测地线具有有限‘长度’”这一要求时，对“长度”一词加了引号。这引号所表示的是对长度定义的推广。具体地说，类时测地线的长度通常是被定义为固有时间的，即

$$\tau = \int ds \qquad (2.1.1)$$

但这一定义不适合描述类光测地线，因为后者对应的固有时间恒为零。因此，为了描述类光测地线，我们需要对长度定义进行推广，推广为所谓的广义仿射参数（generalized affine parameter）。对于一条时空曲线C（t）（t为任意参数），广义仿射参数定义为

$$\lambda = \int \left[\sum_a V^a(t) V^a(t) \right]^{1/2} dt \qquad (2.1.2)$$

其中V^a（t）为曲线在C（t）处的切向量∂/∂t沿该处某标架场e_a（t）的分量（曲线上各点的标架场定义为由某一点的标架场平移而来），求和是欧式空间中的分量求和。显然，这样定义的广义仿射参数是恒为正的，它的数值则与标架场的选择有关。不过可以证明，广义仿射参数的有限与否与标架场的选择无关，从而对于我们表述奇点的定义来说已经足够

了（因为我们只关心“长度”的有限性）。另外值得一提的是，广义仿射参数对所有C^1类（即一次连续可微）的时空曲线都可以定义，而不限于测地线。不难证明，类时测地线的固有时间是广义仿射参数的特例（请读者自行证明）。

作为一个例子，我们来看看施瓦西解中r＝0的奇点是否满足上面所说的奇点定义。为此我们来计算从施瓦西视界（r＝2m）出发，向内（即沿r减小方向）延伸的径向类时测地线的长度（即固有时间）。由施瓦西度规可知

$$ds^2 = -\left(\frac{2m}{r} - 1\right)dt^2 + \left(\frac{2m}{r} - 1\right)^{-1}dr^2 \qquad (2.1.3)$$

因此（请读者补全计算细节）

$$\tau = \int ds < \int\left(\frac{2m}{r} - 1\right)^{-1/2}dr < \pi m < \infty \qquad (2.1.4)$$

由此可见这种测地线的长度是有限的。另一方面，沿这种测地线趋近r＝0时，克莱茨曼标量$R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}$发散，因此这种测地线是不可延拓的。这表明施瓦西解中r＝0的奇点满足上面所说的奇点定义。从物理上讲，这个结果表明落入施瓦西视界的试验粒子会在有限固有时间内从物理时空中消失（形象地说是“落入奇点”）。

现在让我们再回到定义上来，奇点的定义要求时空流形具有测地不完备性。读者也许会问：测地线究竟由于什么原因而不完备？另外，虽说测地不完备性是对时空结构所具有的病态结构的精确描述，但这“精确”二字是以数学上无歧义为标准的。在物理上，我们仍然可以问这样一个问题：当试验粒子沿不完备的测地线运动时，究竟会遇到什么样的时空病态性质？或者简单地说，奇点究竟是什么样子的？对此，物理学家们曾经试图给出直观描述，可惜一直没能找到一种直观描述足以涵盖所有可能的测地不完备性。比如，人们曾经认为奇点的产生意味着某些

几何量（比如克莱茨曼标量）或物理量（比如物质密度）发散，如果是这样，那么沿不完备非类空测地线运动的试验粒子所遇到的将是趋于无穷的潮汐作用或其他发散的物理效应。施瓦西奇点及大爆炸奇点显然都具有这种性质。但细致的研究发现，并非所有奇点都是如此。一个最简单的反例是锥形时空：

$$ds^2 = dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2) \qquad (2.1.5)$$

其中r＞0，0＜φ＜a（a为小于2π的一个角度），并且φ＝0与φ＝a粘连在一起。这个时空是局部平坦的（曲率张量处处为零），并且显然没有任何发散性。但这一时空无法延拓到r＝0（被称为锥形奇点），因而是测地不完备的（类时与类光都不完备）[7]。这个反例表明奇点不一定意味着发散性。

对奇点的另一种直观描述是：奇点是时空中被挖去的点（或点集）。比如施瓦西奇点与刚才提到的锥形奇点是被挖去的r＝0，大爆炸奇点则是被挖去的t＝0。但这种描述如果正确的话，那么通向奇点的所有测地线——无论类时还是类光——必定全都是不完备的。换句话说，如果奇点是时空中被挖去的点（或点集），那么它的存在将同时意味着类时测地不完备性与类光测地不完备性。我们上面举出的所有例子都具有这一特点。但细致的研究表明，这一描述同样不足以涵盖所有的奇点。1968年，美国物理学家杰罗奇（Robert Geroch，1942— ）给出了一个共形于闵科夫斯基时空的时空（R^4，$\Omega^2\eta_{ab}$），其中共形因子Ω^2具有球对称性，在区域r＞1恒为1，在r＝0上满足$t\Omega^2 \to 0$（$t \to \infty$）。显然（请读者自行证明），对于这样的时空，类时测地线r＝0沿$t \to \infty$具有不完备性，因此这个时空流形具有类时测地不完备性。另一方面，所有类光测地线都将穿越区域r≤1而进入平直时空，因而都是测地完备的。由此可见这一时空具有类时测地不完备性，但不具有类光测地不完备性[8]。这个反例表明奇点并非都能理解为是从时空中被挖去的点（或

点集）。

通过这些例子，我们对奇点定义所包含的复杂性有了一些初步了解，它的表述虽然简单，却巧妙地包含了难以完整罗列的种种复杂的时空类型。但另一方面，这个定义虽然已经具有很大的涵盖性，却仍不足以包含所有的奇点类型。这一点也是由杰罗奇指出的，此人在奇点定理的研究中是可以与霍金及彭罗斯齐名的非同小可的人物。1968年，在提出上述反例的同一篇论文中，杰罗奇给出了另外一种时空，它是测地完备的，但却包含长度有限的不可延拓类时曲线（注意是类时曲线而非类时测地线），并且该曲线上的加速度有界。从物理上讲，这意味着在这种时空中，具有有限燃料的火箭所携带的试验粒子沿特定的类时曲线运动，可以在有限时间之内从时空流形中消失。显然，这与自由下落的试验粒子从时空流形中消失具有同样严重的病态性质（事实上这里我们还要多损失一枚火箭！）。因此如果我们认为测地不完备性意味着奇点，那么就必须承认杰罗奇的时空也具有奇点。这个反例表明，我们——以及多数其他文献——所采用的测地不完备性只是定义奇点的充分条件，而不是必要条件。也就是说，一个测地不完备的时空必定具有奇点；但反过来则不然，一个测地完备的时空未必就没有奇点。

物理学家们对奇点性质所做的研究还有许多，限于篇幅，这里不再做进一步介绍，不过在后文介绍宇宙监督假设时我们还会再涉及这一话题。在接下来的几节中，我们将介绍奇点定理及其证明。

2.2　雷查德利方程

在上一节中我们对广义相对论中的奇点作了定义。这样定义的奇点究竟会在什么条件下出现？它是否如某些物理学家猜测的那样来源于对称性？这些都是奇点定理所要回答的问题。

由于我们对奇点的定义是建立在测地不完备性之上的，因此为了研究奇点产生的条件，很自然的做法就是对测地线的性质进行研究。我们用V表示测地线的切矢量，对于类时测地线来说，V满足两个条件：V^aV_a＝1（归一化条件）及$V^aV^b{}_{;a}$＝0（自平移条件，其中“；”为协变导数）。我们效仿线性代数中引进投影算符的做法，引进一个辅助张量h_{ab}＝g_{ab}－V_aV_b。不难证明（请读者自行验证），h^a_b是与V相垂直的子空间上的投影算符，因此h_{ab}有时被称为时空度规g_{ab}的“空间部分”（请读者想一想，这里所说的“空间”是什么含义？）。

我们知道，时空曲率的存在会导致沿相邻测地线运动的试验粒子之间的距离发生变化，这是所谓的测地偏离（geodesic deviation）效应，它是引力相互作用的一种体现。我们对测地线性质的研究也从这个角度入手，考察一个测地线束中的测地偏离效应。对一个测地线束来说，如果我们用与切矢量V相垂直的基矢S表示测地偏离矢量，则两者——作为矢量场——的对易子［S，V］＝0，即（请读者自行证明）：$dS^a/d\tau \equiv V^bS^a{}_{;b} = V^a{}_{;b}S^b$（其中τ为固有时间）。这表明，$V^a{}_{;b}$描述了测地偏离矢量沿测地线的变化。如果我们把沿测地线束运动的一群粒子看成一种类似于连续介质的东西，那么$V^a{}_{;b}$描述的就是这一连续介质的形变。由于这种形变是纯“空间”的（请读者想一想这是什么含义，并且予以证明），因此我们可以仿照连续介质力学的做法，用前面定义的时空度规的“空间部分”h_{ab}将这种形变分解为（请读者加以验证）

$$V_{a;b}=\frac{1}{3}\theta h_{ab}+\sigma_{ab}+\omega_{ab} \qquad (2.2.1)$$

其中θ、σ_{ab}及ω_{ab}分别定义为

$$\begin{cases}\theta=V_{a;b}h^{ab}=V^{a}_{;a}\\ \sigma_{ab}=V_{(a;b)}-\frac{1}{3}\theta h_{ab}\\ \omega_{ab}=V_{[a;b]}\end{cases} \qquad (2.2.2)$$

这里$V_{(a;b)}$与$V_{[a;b]}$分别为$V_{a;b}$的对称与反对称部分。上面这三项均有明确的物理意义：θ被称为膨胀标量（expansion scalar），是$V_{a;b}$的迹，描述的是测地线束会聚或发散的趋势；σ_{ab}被称为切变张量（shear tensor），是$V_{a;b}$的无迹对称部分，描述的是测地线束的空间截面在体积不变（由无迹条件所保证）的情况下产生形变的趋势；ω_{ab}被称为涡旋张量（vorticity tensor），是$V_{a;b}$的反对称部分，描述的是测地线束在空间截面形状不变的情况下相互缠绕的趋势〔9〕。这其中描述测地线束会聚或发散的膨胀标量θ对于奇点定理的讨论有着特别重要的意义，因此我们将着重对它进行研究。

为了研究θ，我们注意到从物理上讲，影响θ的因素是时空曲率（或者说物质分布——两者通过爱因斯坦场方程彼此联系）。因此我们从曲率张量的定义式$V^{a}_{;bc}-V^{a}_{;cb}=R^{a}_{dbc}V^{d}$出发〔10〕。将这一表达式对指标a和b进行缩并，与V^c取内积，并利用$V_{a;b}$的分解式及类时切向量V的性质，便可证明θ沿测地线的变化为

$$\frac{d\theta}{d\tau}\equiv V^{a}\theta_{;a}=-R_{ab}V^{a}V^{b}-\frac{1}{3}\theta^{2}-\sigma_{ab}\sigma^{ab}+\omega_{ab}\omega^{ab} \qquad (2.2.3)$$

其中τ为固有时间。这个方程被称为雷查德利方程（Raychaudhuri

equation）[11]，是印度物理学家雷查德利（Amal Raychaudhuri，1923—2005）与俄国物理学家朗道（Lev Landau，1908—1968）彼此独立地提出的。雷查德利方程的提出恰好是在爱因斯坦逝世的那一年（1955年），它与能量条件的结合将成为证明奇点定理的重要环节。

2.3　测地线束与共轭点

在雷查德利方程中，如果所考虑的测地线束局部正比于某个梯度场，或者说垂直于某个超曲面，则称该线束是超曲面垂直（hypersurface orthogonal）的。可以证明，对于这样的测地线束来说，涡旋张量ω_{ab}为零，从而雷查德利方程可以简化为

$$\frac{d\theta}{d\tau} = -R_{ab}V^aV^b - \frac{1}{3}\theta^2 - \sigma_{ab}\sigma^{ab} \qquad (2.3.1)$$

由于$\sigma_{ab}\sigma^{ab}$总是非负的，因此从这个方程中我们可以得到

$$\frac{d\theta}{d\tau} \leqslant -R_{ab}V^aV^b - \frac{1}{3}\theta^2 \qquad (2.3.2)$$

如果进一步假定强能量条件成立，即$R_{ab}V^aV^b$处处非负，则上述不等式可以进一步简化为

$$\frac{d\theta}{d\tau} \leqslant -\frac{1}{3}\theta^2 \qquad (2.3.3)$$

对这个不等式进行积分可得

$$\theta^{-1} \geqslant \theta_0^{-1} + \frac{1}{3}(\tau - \tau_0) \qquad (2.3.4)$$

其中$\theta_0=\theta(\tau_0)$。从这个不等式我们可以得到一个重要的推论，那就是倘若$\theta_0<0$，即线束在$\tau=\tau_0$时出现会聚效应，则θ会在有限固有时间$\tau-\tau_0\leq 3/|\theta_0|$内趋于负无穷。可以证明，这意味着测地线束在该处会聚为一点，或者说测地偏离矢量场——也称为雅可比场（Jacobi　field）——在该处为零。如果一个从p点发出的非平凡（即各测地线不处处重合，或者说雅可比场不处处为零）的类时测地线束在q点会聚，我们就把q和p

称为该测地线束上（即其中每一条测地线上）的一对共轭点（conjugate points）。从上面的分析中我们看到，如果从p点发出的一个类时测地线束在未来某一点上出现会聚效应$\theta<0$，则在该线束上距离p有限远的地方必定存在一个与p共轭的点q——当然，这里我们要假定该测地线束可以延伸到q点。

显然，在一个测地完备时空中，“测地线束可以延伸到q点”这一假定是自动满足的。因此，对于测地完备时空来说，上面这个结果是所有类时测地线都满足的普遍性质。进一步的分析表明，上述结果所要求的条件，即测地线束在“某一点上出现会聚效应$\theta<0$”，可以转化为一个有关曲率张量的条件。事实上，由式（2.3.1）和式（2.3.4）可以看到，即便在$\sigma_{ab}\sigma^{ab}$与$R_{ab}V^aV^b$处处为零（此时式（2.3.4）取等号形式），且$\theta_0>0$这一对于形成$\theta<0$来说最为不利的条件下，θ仍将在$t\rightarrow\infty$时趋于零（即几乎就要形成$\theta<0$这一结果）。这使人想到，上述最为不利的条件只要在某一点上（从而由连续性条件可知在该点的一个邻域内）被破坏，比如$R_{ab}V^aV^b>0$在某一点上成立，就足可造成当τ足够大时$\theta<0$。可以证明，事实的确如此。因此“某一点上出现会聚效应$\theta<0$”这一条件可以转化为某一点上$R_{ab}V^aV^b>0$。如果我们进一步把$\sigma_{ab}\sigma^{ab}$所起的作用也考虑进去，这一条件还可以继续减弱，最终可以得到这样一个结果：在一个测地完备的时空中，如果强能量条件成立，并且在每条类时测地线上至少有一个点使得$R_{abcd}V^bV^d\neq0$，则所有类时测地线上都存在共轭点对，简称共轭对。

从物理意义上讲，每条类时测地线上至少有一个点使得$R_{abcd}V^bV^d\neq0$，意味着每条类时测地线都至少会在一个时空点上遇到由物质分布或引力波所造成的某种测地偏离效应。这一条件——称为类时一般性条件（timelike generic condition）——在理论上可以被一些非常特殊的情形（比如曲率张量与测地线切矢量形成特殊分量匹配的情形）所违反。但对于具有现实物理意义的情形来说，由于物质及引力波的分

布往往足够弥散及随机，类时一般性条件被认为是得到满足的。

上面这些结果都是针对类时测地线的。不过可以证明，除了一些不影响定性结果的差异（比如固有时间τ变成仿射参数λ，雷查德利方程中的数值因子1/3因垂直子空间维数的改变而变成1/2等）外，类光测地线也具有类似性质。类光测地线所满足的一般性条件为：每条类光测地线上至少有一个点使得$k_{[e}R_{a]bc[d}k_{f]}k^b k^c \neq 0$。这个条件被称为类光一般性条件（null generic condition）〔12〕。类时与类光一般性条件统称为一般性条件〔13〕。把类时与类光情形合在一起，我们前面所介绍的结果可以重新表述为：在一个测地完备的时空中，如果强能量条件与一般性条件成立，则每条非类空测地线上都存在共轭对〔14〕。这是一个不依赖于对称性的普遍结果，它对于奇点定理的证明及确立奇点定理的普适性都有极其重要的作用。

细心的读者可能还记得，在上述结果的证明伊始，我们曾经作过一个假设，即所考虑的测地线束是超曲面垂直的。这个假定保证了ω_{ab}＝0，从而消除了雷查德利方程中与其他各项符号相反——因而会对我们的证明造成极大干扰——的$\omega_{ab}\omega^{ab}$项（请读者想一下，这一项的符号与其他各项相反的物理意义是什么？）。那么这个假设具有多大的普遍性呢？或者说，这个假设是否会使上述结果——进而使整个奇点定理的证明——失去应有的普遍性呢？答案是否定的，因为在数学上可以证明，经过某一时空点的类时测地线束必定在该点的某个凸邻域内具有超曲面垂直性，因此ω_{ab}在该邻域内必定为零。不仅如此，通过一个与雷查德利方程类似的描述ω_{ab}沿测地线变化的方程可以证明，ω_{ab}沿一条测地线只要在某一点上为零，就沿该测地线处处为零。因此，假定测地线束为超曲面垂直不会有损结果的普遍性。

综合上述分析，我们看到，在一个具有适当物质分布的测地完备时空中共轭点的存在是普遍现象。假如有一个适当的物质粒子群沿某个非

类空测地线束运动，那么当它们运动到共轭点上时，由于测地线的会聚，粒子的数密度（以及质量密度）将趋于发散，从而形成一个奇点。雷查德利发表于1955年的原始论文就涉及了这样的情形[15]。不过，在一般情况下显然并没有理由假定存在那样的物质粒子群，因此共轭点的存在只是一个抽象的几何结果，不会直接导致奇点，上述结果也不足以作为奇点存在性的证明（如果一定要算证明的话，只能算是非常弱的证明，因为它所要求的条件太过特殊）。但是，这一结果为十年后彭罗斯等人的工作奠定了基础，是证明奇点定理的第一步。这一步所侧重的是引力理论中的动力学因素，强能量条件的引进是这种因素的体现。

在下一节中将会看到，当我们把有关测地线的上述结果与看似风马牛不相及的时空的因果性质结合起来时，奇点在广义相对论中的出现就变得几乎不可避免了。

2.4　时空的因果结构

在前两节中，我们介绍了证明奇点定理的第一步，即通过雷查德利方程、强能量条件及一般性条件，确立了测地完备时空中每条非类空测地线上都存在共轭对这一结论。那一步侧重的是引力理论中的动力学因素。接下来的第二步侧重的则是时空的因果结构。

在物理学中，因果性是一个很微妙的概念。一方面，它是现实世界中最基本的经验事实之一；另一方面，却很少有物理理论直接把因果性作为前提条件。由此导致的一个结果是：某些物理理论起码在形式上允许因果性的破坏。广义相对论就是这样的一个理论。在广义相对论中，破坏因果性最简单的方式是产生闭合非类空曲线（请读者想一想，这种曲线在什么意义上破坏因果性？）。为了对这种类型的因果性破坏进行界定，人们引进了一个条件，叫做因果性条件（causality condition）：一个时空如果不存在闭合非类空曲线，则称为满足因果性条件。考虑到所有有质量粒子都只能沿类时曲线运动，人们还提出了一个比因果性条件稍弱的条件，称为时序条件（chronology condition），它与因果性条件的差别在于把“不存在闭合非类空曲线”减弱为“不存在闭合类时曲线”。

虽然时序条件比因果性条件稍弱，但可以证明，一个时空如果在时序条件之外还满足测地完备性，则该时空将不仅满足因果性条件（即不存在闭合非类空曲线），而且不存在可以无限逼近闭合非类空曲线的曲线。这个比因果性条件更强的性质被称为强因果性条件（strong causality condition），它在奇点定理的研究中是一个重要概念。

奇点定理研究中的另一个重要概念是所谓的封闭陷获面（closed trapped surface）。这是一种特殊的二维封闭类空曲面，其基本性质是：所有与之正交的类光测地线束无论向内还是向外都是趋于会聚的（即膨

胀标量θ＜0）。从物理上讲，该性质意味着从封闭陷获面发出的光波的波前是收缩的。这种曲面在广义相对论中并不鲜见，比如施瓦西解中所有r＜2m的曲面都具有这一性质（这表明任何物质——包括光波——都不能从施瓦西黑洞中逃脱）。1983年，美国数学家舍恩（Richard Schoen，1950—　）与美籍华裔数学家丘成桐（Shing-Tung　Yau，1949—　）证明了一个相当普遍的结果：只要物质的分布足够致密，就必定会出现封闭陷获面。

由于封闭陷获面的定义是建立在类光测地线的行为之上的，因此我们引进与类光测地线有关的两个特殊点集：E^+（S）与E^-（S），分别由从曲面S发出的未来与过去方向的类光测地线所组成〔16〕。在这一定义中我们还进一步假定S上任意两点之间都不存在类时连接，这种点集S被称为非时序点集（achronal　set）。封闭陷获面由于是类空的，显然也是非时序点集。可以证明，如果强能量条件成立，则对于任何封闭陷获面S，E^+（S）与E^-（S）紧致。

我们知道，物理上所有的相互作用都是非类空传播的（也就是说相互作用的传播速度不大于光速）。因此，如果我们考虑时空中某一点上的任何物理性质，它所能依赖的初始条件只能位于与该点具有非类空连接的时空点上。反过来说，给定某个时空区域S上的初始条件，我们能完全确定其性质的时空区域是由那样的一些点组成的：所有通过那些点的过去不可延拓非类空曲线都与S相交。这一时空区域被称为S的未来柯西展开（future　Cauchy　development）或未来影响域（future　domain　of　dependence），通常记为D^+（S）。D^+（S）的边界则被称为未来柯西视界（future Cauchy horizon），记为H^+（S）。类似地，我们也可以定义S的过去柯西展开（或过去影响域）和过去柯西视界，分别记为D^-（S）和H^-（S）。S的未来柯西展开与过去柯西展开合在一起——即D^+（S）∪D^-（S）——称为S的柯西展开（或影响域），记为D（S）。一个时空（或时空中的一个点集）M中如果存在一个封闭非时

序点集S，使得M＝D（S），则称为是全局双曲（globally hyperbolic）的，相应的封闭非时序点集S（可以证明它一定是一个超曲面）被称为柯西面（Cauchy surface）。柯西面可以被形象地理解为时空中对应于某一时刻的超曲面。一个时空如果是全局双曲的，我们就可以通过柯西面上的初始条件预言整个时空中的演化，因此时空的全局双曲是一种非常优良的因果性质。1965年，彭罗斯正是在假设时空为全局双曲的基础上证明了最早的奇点定理。但是，时空的全局双曲是一个很强的假设，要想证明现实时空满足这样的假设几乎是不可能的。因此五年之后（即1970年），霍金与彭罗斯放弃了这一假设，在一组物理上更容易实现的假设之上重新证明了奇点定理，这便是我们所要介绍的霍金-彭罗斯奇点定理。

对于奇点定理的证明来说，全局双曲时空（或点集）有一个很重要的性质，那就是其中任意两个可以建立非类空连接的时空点p和q之间必定存在一条非类空测地线，其长度最大——即大于或等于p和q之间的任何其他非类空曲线（请读者想一想，为什么这里测地线的长度是最大而不是最小？），并且在p和q之间不存在与p共轭的点。

2.5 霍金-彭罗斯奇点定理

以上我们介绍了一些在奇点定理的研究中常用的有关时空因果性质的定义及结果。虽然介绍得比较零散，但有些读者可能已经看出一点思路来了：我们在2.3节中曾经证明了，在适当的条件下，每条非类空测地线上都存在共轭对；而在上节的末尾我们则开始接近一个与之矛盾的结果，即在特定的条件下，某些测地线上不存在共轭对。这对彼此矛盾的结果正是证明奇点定理的关键。确切地讲，奇点定理的证明是要通过这对彼此矛盾的结果来论证以下5个条件不可能同时成立：

（1）时空是测地完备的；

（2）强能量条件成立；

（3）一般性条件成立；

（4）时空满足时序条件；

（5）时空中存在一个非时序点集S，使得E^+（S）与E^-（S）紧致。

限于篇幅，我们只能简单叙述一下论证思路。在上述5个条件中，条件（1）～（3）是2.3节所介绍的证明奇点定理的第一步中所用到的条件，由此推知的是每条非类空测地线上都存在共轭对。条件（1）和条件（4）所推知的——如上文所述——是时空满足强因果条件。而由强因果条件与条件（5）则可以证明这样一个结果：时空中存在一个全局双曲区域M，其中包含一条未来不可延拓类时曲线γ及一条过去不可延拓类时曲线λ。利用这一结果就可以证明时空中存在一条没有共轭对的非类空测地线。具体做法是：在λ上取一个沿过去方向趋于无穷的点集a_n，同时在γ上取一个沿未来方向趋于无穷的点集b_n（选取时使得b_1在a_1的类时未来，从而保证所有b_n都在a_n的类时未来）。由于M是全局双曲的，因此——由上节末尾的结果可知——在每一对a_n和b_n之间都存在

一条（长度最大的）非类空测地线μ_n，其上在a_n和b_n之间不存在a_n的共轭点。可以证明，存在于M之中的这一由非类空测地线μ_n组成的无穷集合必定有一个聚点（limit point）μ，它是一条非类空测地线，并且其上不存在任何共轭对。这样，我们就得到了与第一步（也就是条件（1）～（3））所得的“每条非类空测地线上都存在共轭对”相矛盾的结论，从而证明了上述5个条件不可能同时成立。

既然上述5个条件不可能同时成立，那么我们就可以选其中4个条件为前提（即假定这4个条件成立），来推翻剩下的那个条件〔17〕。霍金与彭罗斯所做的是以条件（2）～（5）为前提，来推翻条件（1），即证明时空不是测地完备的。按照我们在2.1节所作的定义，这表明时空中存在奇点。这就是霍金与彭罗斯的奇点定理。

在被奇点定理采用为前提的条件（2）～（5）中，条件（2）～（4）都有明确的物理意义，唯独条件（5）——即时空中存在一个非时序点集S，使得E^+（S）与E^-（S）紧致——显得很抽象。幸运的是，我们可以用一些物理意义更为明确的条件来取代这一抽象的数学条件。在上节中我们介绍过，如果强能量条件成立，则对于任何封闭陷获面S，E^+（S）与E^-（S）紧致。由于强能量条件已经包含在条件（2）～（4）中了（即条件（2）），因此我们可以用“时空中存在封闭陷获面”来替代条件（5），这个条件在物理上可以由足够致密的星体来满足。除此之外，霍金与彭罗斯还提出了另外两个条件来替代条件（5）：一个是“时空中存在紧致无边的非时序点集”〔18〕，这个条件在物理上可以由空间上有限无边的宇宙来满足；另一个是“时空中存在一个点，通过该点的所有未来（或过去）方向的类光测地线束的膨胀标量θ最终将变为负值”，这个条件在物理上可以由局部膨胀或收缩的宇宙来满足。这3个替代条件都被认为是原则上可以检验，并且很可能在我们的宇宙中已经得到满足的条件。

至此，我们可以对霍金与彭罗斯所证明的奇点定理做一个完整表

述：

霍金-彭罗斯奇点定理：一个时空若满足以下条件，就必定是非类空测地不完备的（即存在奇点）：

（1）强能量条件成立。

（2）一般性条件成立。

（3）满足时序条件。

（4）以下三个条件之一成立：

①存在封闭陷获面；

②存在紧致无边非时序点集；

③存在一个点，通过该点的所有未来（或过去）方向的类光测地线束的膨胀标量θ最终将变为负值。

这个定理是霍金与彭罗斯于1970年提出并证明的。如我们在上文中所说，这并不是最早的奇点定理。彭罗斯于1965年、杰罗奇于1966年、霍金于1967年都提出过奇点定理。比较之下，霍金-彭罗斯奇点定理所要求的条件在物理上最容易实现，并且涵盖面也广〔19〕，因此人们如今提到奇点定理时通常指的就是这一定理。霍金-彭罗斯奇点定理不依赖于对称性，它对于确立广义相对论中奇点的存在性及普遍性是非常重要的。同时它也是对我们在1.1节中所介绍的有关奇点的不同见解的有力裁决。但是，霍金-彭罗斯奇点定理也有一个显而易见的缺点，那就是它既无法告诉我们究竟哪一条非类空测地线是不完备的，也无法提供有关奇点具体性质的信息。这一缺点为后人加强奇点定理的结论部分留下了空间。不过要想加强奇点定理的结论部分，往往不可避免地要对前提部分也予以加强，从而有损定理的普遍性。

2.6　讨论

我们对奇点定理的介绍就要结束了。有些读者可能会提出这样一个问题：我们证明霍金-彭罗斯奇点定理所用的是排除法，即通过证明测地完备性与奇点定理的4个前提不相容，来排除测地完备性，从而确立奇点的存在。但是，当一组命题不相容时，究竟哪个（或哪几个）命题应该被排除，在逻辑上是有很大随意性的[20]。因此从逻辑上讲，由上面介绍的不相容性，原则上可以通过排除不同的命题，而得到不同的定理（读者不妨自己写出几个看看）。为什么我们偏偏要选择将时空的测地完备性作为被排除的命题，从而得到霍金-彭罗斯奇点定理呢？

这是一个非常好的问题。我们知道，一个物理上有价值的定理必须能对物理世界作出某种程度的描述。因此，在所有逻辑上成立，并且能进行物理诠释的数学命题中，只有那些其前提在物理上能够实现的定理才能成为有效的物理定理。如果已经知道物理世界不满足某一性质，那么把该性质作为前提的数学命题就不能成为有效的物理定理。从这个意义上讲，我们可以通过考察霍金-彭罗斯奇点定理所涉及的4个前提在物理世界中实现的可能性，来分析这一定理的合理性。

在霍金-彭罗斯奇点定理的4个前提中，前提（4）属于初始及边界条件，并且实现的可能性极大。事实上，早在霍金-彭罗斯奇点定理提出的年代，天文观测及理论研究就已经在很大程度上显示出这个前提的3个子条件很可能部分甚至全部得到满足。前提（1）和前提（2）与人们在宏观世界的观测经验相符，因为迄今所知的所有宏观物质的能量动量张量都满足强能量条件，而现实宇宙中物质（包括宇宙微波背景辐射）及引力波的分布无疑遍及全空间，从而满足一般性条件，因此在以大尺度宏观世界为主要描述对象的广义相对论中，这两个前提被认为是适用的。前提（3）所要求的不存在闭合类时曲线也具有不错的经验基

础，因为时间的单向性是宏观世界中最基本的经验事实之一。因此所有这4个前提都有可信赖之处，这在很大程度上保证了霍金-彭罗斯奇点定理的价值。

但如果一定要在这4个前提中找出一个最有可能在现实物理世界中不成立的，那么——如我们将在后文中看到的——能量条件（即前提（1））将是首选，因为理论与观测都表明它事实上就不成立。不过，能量条件的破坏主要来自量子效应，而我们所讨论的奇点定理是经典广义相对论中的命题，两者在所涉范围上是有出入的。那么，假如我们不考虑量子效应，或者说只考虑经典广义相对论，又有哪一个前提最值得怀疑呢？一般认为是时序条件（即前提（3））。这一条件要求不存在闭合类时曲线。它之所以值得怀疑，主要有两个原因：一是因为广义相对论的某些特殊解事实上允许闭合类时曲线存在，虽然迄今为止那些解还没有一个得到过任何观测上的支持；二是由于闭合类时曲线实际上是一种抽象的时间机器，这是一种在很多方面都很引人入胜的东西。因此有些物理学家把广义相对论没有在原理层面上禁止闭合类时曲线，视为是一个很值得探索的理论问题。

如果时序条件有可能被破坏，那就产生了一个很自然的问题：我们是否可以通过作一个与霍金-彭罗斯奇点定理不同的选择，把测地完备性作为定理的前提之一，而把时序条件的破坏（从而允许时空中存在闭合类时曲线）作为定理的结论呢？[21]对这种可能性物理学家们也进行过一些研究。1977年，美国图兰大学（Tulane University）的物理学家梯普勒（Frank Tipler，1947—　）研究了渐近平直时空中有限大小的闭合类时曲线，结果发现在强能量条件与一般性条件等成立的情况下，这样的曲线在测地完备时空中是不可能出现的[22]。其他一些物理学家后来也做了这方面的研究和推广，包括使用更弱的条件，以及推广时序破坏的定义等，得到的结果都类似。这些结果成为后来霍金提出所谓时序保护假设（chronology protection conjecture）的基础之一。这些结

果表明，时序条件的破坏在很大程度上本身就意味着测地完备性的破坏，因而放弃时序条件并不能挽回测地完备性〔23〕。这在一定程度上进一步加强了奇点的不可避免性〔24〕，也进一步支持了霍金-彭罗斯奇点定理的合理性——当然，所有这一切都限于经典广义相对论的范围。

2.7 附录：雷查德利小传

2005年6月18日，印度物理学家雷查德利（见图2.1）因心脏病发作猝然去世，享年81岁。雷查德利是为数不多的名字进入科学术语的印度物理学家之一，但他的名字对于多数人——包括他的印度同胞以及世界各地广义相对论专业以外的物理系学生——来说都显得比较陌生，他的生平更是鲜为人知。在这里，我们将对雷查德利的生平作一个简单介绍，作为奇点与奇点定理介绍的背景资料。同时笔者也希望本节能对以著名科学家为主要关注对象的中文传记资料起到一定的补充作用。从某种意义上讲，了解一些像雷查德利这样的“非著名科学家”的生平，对于今后从事科学研究的学生来说，或许会有一些从名人传记中无法得到的启示——因为他们的人生和际遇更接近常人。



图**2.1**　印度物理学家雷查德利（**1923—2005**）

雷查德利于1923年9月14日出生在巴里萨尔（Barisal），这是当时英属印度（British India）的一部分，如今则是孟加拉国中南部的一个货物转运中心。雷查德利的父亲是印度东北部大城市加尔各答（Calcutta）一所初级学校的数学教师，这座城市后来也成为雷查德利

一生主要的学习和工作地点。或许是受父亲的影响，雷查德利自小就对数学怀有兴趣，并且成绩优异。

中学毕业后，雷查德利进入了印度著名学府加尔各答大学（University of Calcutta）附属的院长学院（Presidency College）。他当时最感兴趣的是数学，其次则是物理。不过在选择专业这个节骨眼上，雷查德利的父亲再次对他产生了影响。在父亲的建议下，物理成为了雷查德利的专业。一位数学教师有一位对数学感兴趣的儿子，他为什么不让儿子“子承父业”呢？这是因为雷查德利的父亲虽只是一名不起眼的初级学校教师，却有着当时极为出众的数学硕士学位，他的这种大材小用的郁闷经历使得他希望自己的儿子能尝试一个不同的专业。

1942年，雷查德利顺利地获得了学士学位，两年后又获得了加尔各答大学的硕士学位，平了父亲的学位纪录。1945年，雷查德利成为印度科学培训协会（India Association for the Cultivation of Science，IACS）的一名研究人员，开始了自己的学术生涯。印度科学培训协会虽然有着很土的名字，却是印度历史最悠久的基础科学研究院。印度最著名的本土物理学家拉曼（Chandrasekhara Raman，1888—1970）正是在这里发现了以他名字命名的拉曼效应（Raman effect），并获得1930年的诺贝尔物理学奖。但对雷查德利来说，进入声名卓著的科学培训协会并未使自己的学术之路成为坦途。那时候，他对广义相对论产生了兴趣，但科学培训协会却要求他从事实验工作，因为广义相对论在当时已被视为是很难做出新成果的学科。由于谋职不易，雷查德利痛苦地服从了分配，但他凭借自己的兴趣自学了微分几何与广义相对论。四年后，雷查德利在加尔各答大学附属的阿属托什学院（Asutosh College）获得了一个临时教职。在那期间，他对广义相对论中施瓦西解的奇点进行了研究，这一研究在很大程度上预示了他一生的兴趣〔25〕。

雷查德利在广义相对论方面的研究没能使他在阿属托什学院的临时职位变为永久。1952年，他重新回到了印度科学培训协会。距他上一次

进入该协会，时间虽已相隔七年，但广义相对论不被看好的局面依然如故。雷查德利此次被要求从事的是金属性质研究。与在阿属托什学院的职位一样，雷查德利在科学培训协会的职位也是临时的，需要发表一定数量的论文来维持。不过，在这种被迫在一个自己不感兴趣的领域从事工作的压力干扰下，雷查德利并未放弃对广义相对论的研究。1955年，他在一篇研究奇点问题的文章中提出了如今以他名字命名的雷查德利方程〔26〕。他的这一工作为十年后彭罗斯与霍金等人证明奇点定理做了重要铺垫，也为他所热爱的广义相对论重新成为研究热点起到了间接作用〔27〕。不过，雷查德利的这一工作在印度并未引起反响，最早的支持来自于遥远的欧洲，据说量子力学创始人之一的约尔当（Pascual Jordan，1902—1980）曾在一个讨论会上称赞雷查德利的工作。比雷查德利稍晚，著名俄国物理学家朗道也独立地提出了雷查德利方程。此后几年，雷查德利方程的重要性渐渐受到一些广义相对论学者的注意，这一成功给了雷查德利很大的信心和勇气，他以此为基础撰写了博士论文。雷查德利的博士论文受到了著名广义相对论学者惠勒（John Wheeler，1911—2008）的好评。1959年，他获得了物理学博士学位〔28〕。

1961年，雷查德利的母校院长学院为他提供了一个理论物理教授的职位，雷查德利回到了阔别二十年的熟悉校园，那里从此成为了他的学术归宿。在此后的漫长岁月里，雷查德利曾几度申请研究所的职位，但都没有成功。在学术生涯的主要时间里，他几乎没有获得任何荣誉，以至于一些印度物理学家在他去世后沉痛地将他称为最受忽视的印度物理学家。雷查德利在院长学院期间为印度的物理教育做出了卓越贡献。印度独立后，为促进科研，追赶西方国家在科技领域的领先地位而建立了不少纯研究性的机构。那些机构具有非常独立的地位，设备与条件均优于大学，因而吸引了大量的研究人员。很多年后，印度科学界开始反思那些机构对印度高等教育的负面影响。简单的说，那些机构对研究人员

的吸积作用，以及它们与高校完全脱钩的体制，使得印度第一流的研究人员脱离了学生，而第一流的学生又找不到好的教师。在那样的环境下，雷查德利默默无闻地“陷落”在高校，反倒成为了印度物理学界的幸事。雷查德利的一些学生后来成为了印度重要的物理学家，其中包括像森（Ashoke Sen，1956— ）这样具有国际知名度的理论物理学家。雷查德利去世后的第二天，他的一位学生在给同事的电子邮件中写道：“我们中有幸在本科阶段成为他学生的许多人，都把他视为我们一生遇到的最伟大的物理老师。”

1972年，印度天体物理学家纳里卡（Jayant Narlikar，1938— ）从英国剑桥“海归”到了印度的塔塔基础研究所（Tata Institute of Fundamental Research）。他的回国不仅将一些前沿的物理信息带回了印度，也把雷查德利方程在西方国家广义相对论研究中所起的作用反馈回了印度。雷查德利的声望终于开始上升，渐渐成为印度一位受人尊敬的物理学家。印度物理学界忽视了雷查德利那么久，但最终，在雷查德利晚年以及去世之后，毕竟还是他们怀念和记录了他的生平。而西方国家的广义相对论学者虽然对雷查德利方程给予了很大的重视，对雷查德利本人的了解却微乎其微。在霍金等人的广义相对论专著中都介绍并使用了雷查德利方程，但都没有直接提及雷查德利的原始论文，更不用说对他进行片言只语的生平介绍了。作为物理学家的雷查德利完全湮没在了作为方程名称的雷查德利的背后。我们现在所知有关雷查德利的生平资料，几乎全都出自印度物理学家之手（当然，其中不免有一些过誉之语）。在雷查德利去世前，印度校际天文与天体物理中心（Inter-University Centre for Astronomy and Astrophysics）及印度科技部所设的科普组织“科学传播”（Vigyan Prasar——这个英文名称是印地语音译）拍摄了一部以他的生平和成就为内容的纪录片。可惜雷查德利本人没能看到这部纪录片的问世。

雷查德利去世的时间距离他发表雷查德利方程恰好相隔半个世纪。

2004年2月，雷查德利在《广义相对论与引力》（General Relativity and Gravitation）杂志上发表了一篇题为“非旋转理想流体的无奇点宇宙学解”（Singularity-Free Cosmological Solutions with Non-Rotating Perfect Fluids）的文章。这是他一生所写的最后一篇论文，对奇点及相关课题的研究伴随他走完了整个孤独的学术生涯。

注释

〔1〕当然，这里所谓的“其他物理理论”指的是不把时空本身作为研究对象的理论。另外，人们通常把由非奇异初始条件演化而来的奇点称为理论本身所具有的奇点，以区别于纯粹通过初始条件而引入的奇点。显然，线性理论——比如麦克斯韦电磁理论——不可能有这样的奇点。

〔2〕洛伦兹度规是指符号差（signature）为（1，—1，—1，—1）的度规（有些文献的定义与我们的相差一个整体符号）。除洛伦兹度规外，人们还常常在时空定义中附加一些其他条件，比如豪斯道夫条件（Hausdorff condition）、连通性，等等。对于度规的可微性则有的假定为C^{∞}，有的假定为C^{r}（r为正整数——请读者思考一下，r最小应该是多少？），等等。

〔3〕有些物理学家试图将奇点视为时空流形的边界，称为奇异边界（singular boundary），但这方面迄今尚未建立起令人满意的处理方式。

〔4〕非类空是类时与类光的统称。这里所说的试验粒子包括零质量粒子。

〔5〕一条时空曲线不可延拓，直观地说就是在时空流形内没有端点。为此我们首先要求时空流形本身是“不可延拓”的，即无法等度规地（isometrically）嵌入更大的流形中（不过如何实现这一要求本身就是一个很艰深的问题）。这一要求排除了一些平凡的奇点，比如在闵科夫斯基时空（Minkowski spacetime）中挖去一个时空点所造成的“奇点”。测地线的不可延拓性则可以用来排除诸如施瓦西视界这样的表观奇点。

〔6〕我们也可以完全类似地定义类空测地不完备性，但由于沿类空测地线的运动是物理上不可实现的，因此这种测地不完备性在奇点研究中不如其他两种测地不完备性那样受重视。

〔7〕这个例子比较平凡，一个更复杂的例子是所谓的Taub-NUT空间，它具有$R^1 \times S^3$拓扑结构，曲率张量处处有界，但同样是测地不完备的（类时与类光都不完备）。

〔8〕这个例子比较特殊，一个更具物理意义的例子是雷斯勒-诺斯特朗姆（Reissner-Nordström）解，它描述的是带质量及电荷的球对称时空，雷斯勒-诺斯特朗姆解具有类光测地完备性，但不具有类时测地不完备性。

〔9〕文献中对这一张量的称呼很多，除“涡旋”外，常见的叫法还有“扭变”（twist）与“旋转”（rotation）。

〔10〕确切地讲，这是将曲率张量的定义用于测地切矢量场所得到的关系式。

〔11〕需要提醒读者注意的是，这里介绍的只是雷查德利方程应用于测地切矢量场的特例，一般情况下雷查德利方程的适用范围并不限于测地切矢量场。

〔12〕类时一般性条件也可以表示成这一形式，因此有些文献对两类一般性条件统一使用这一形式。

〔13〕从某种意义上讲，一般性条件也是对物质分布的间接限定。这种限定不同于我们在1.2节中介绍的那些能量条件，因为平直时空满足所有那些能量条件，却不满足一般性条件。有鉴于此，有些文献把一般性条件通俗地表述为“时空中存在物质”。不过要注意的是：由于一般性条件涉及的是曲率张量而不是里奇（Ricci）张量，因此不能被简单地约化为对物质能量动量张量的直接限制。

〔14〕在某些文献中，强能量条件与一般性条件被合称为一般能量条件（generic energy condition）。

〔15〕雷查德利当时研究的是非旋转尘埃物质的运动。在雷查德利之前，奥地利逻辑学家哥德尔（Kurt Gödel，1906—1978）在1949年发表的有关哥德尔解的著名论文中率先用到了一些在奇点定理的证明中有着重要作用的技巧及思路。这些工作都是奇点定理的先声。

〔16〕E^{+}（S）和E^{-}（S）分别被称为S的未来边界（future horismos）和过去边界（past horismos）。

〔17〕严格讲，我们还必须证明被选为前提的那4个条件彼此相容（因为5个条件不相容并不保证其中4个一定相容）。对于在奇点定理中假定为前提的条件（2）～（5）来说，这一证明可以通过给出一个满足条件（2）～（5）的广义相对论的解来完成。

〔18〕这种非时序点集S的边界定义为这样一些点p的集合：p的任何开邻域都包含一个位于p的类时过去的点，一个位于p的类时未来的点，以及一条连接这两个点但与S不相交的类时曲

线。可以证明，这种非时序点集必定是三维子流形。

〔19〕在本系列的讨论中我们没有对时空流形的微分结构进行细分。如果细分的话，霍金-彭罗斯奇点定理要求度规张量起码是二阶连续可微（即C^2）的。这一条件对于某些严格解——比如物质密度存在突变的解——来说并不成立。不过对于具有现实物理意义的情形来说，度规张量即物质密度通常都被认为是足够可微的。

〔20〕举一个最简单的例子：由$\neg$（A∧B）（即A和B不能同时成立）既可以推知A→$\neg$B（即由A成立推出B不成立），也可以推知B→$\neg$A（即由B成立推出A不成立）。

〔21〕当然，这样的定理将不再是奇点定理，而应该被称为时序破坏定理，或时间机器存在定理。

〔22〕在这一研究中，梯普勒还假定了物质的能量密度在过去类时曲线上处处大于一个非零正值，所得到的结果则是时空必定是类光测地不完备的（从而也是测地不完备的）。

〔23〕有读者可能会问：既然无论时序条件是否被破坏，奇点都会出现，那为什么不干脆把时序条件从奇点定理的前提中去掉呢？这是因为在论证时序条件的破坏会导致测地完备性的破坏时，往往要引进一些额外的条件（比如前一个注释中提到的梯普勒所用的额外条件）。这些看似细微的额外条件的使用，使得我们无法将时序条件从奇点定理的前提中简单地去除。

〔24〕奇点的不可避免性也受到诸如施瓦西解等物理上相当合理的解的直接支持。不过需要指出的是，从纯理论的角度讲，奇点定理所要求的许多条件都是可以被突破的，人们也的确因此而构造出了没有奇点的解。只不过那些解都是非常特殊的，对于具有现实意义的解，广义相对论中奇点的出现几乎是不可避免的。

〔25〕当时包括爱因斯坦在内的许多人认为施瓦西解在r＝2m处的“奇点”是不可到达的，雷查德利的论文对这种看法提出了异议。雷查德利去世后，一些印度物理学家对这一工作给予了很高的评价。不过那些评价有些过誉，因为r＝2m是否可以到达之所以成为问题，在很大程度上是由于当时许多人误将r＝2m当成奇点（即所谓的“施瓦西奇点”）。而r＝2m的非奇异性其实早在1933年就由勒梅特等人指出过（参阅1.1节的注释），那才是解决“施瓦西奇点”问题的关键性工作。

〔26〕这一工作其实1953年就完成了，花了两年时间才得以发表。

〔27〕有读者可能会问：雷查德利比彭罗斯等人早了整整十年就对奇点进行研究，并且得

到了在奇点定理的证明中起到重要作用的雷查德利方程，为何却没能率先证明奇点定理？这其中主要的原因是因为在奇点定理的证明中还需要大量严密而精巧的几何论证，以及对时空因果性质的细致分析。从雷查德利的研究风格来看，那样的论证和分析明显超出了他的数学背景和研究范围。另一方面，雷查德利一直在试图寻找的没有奇点的时空，这种动机也在无形中影响了他进一步努力的方向。

〔28〕在科学培训协会期间，雷查德利在其“正业”金属性质研究中也取得了一定的成果。

第3章
正质量定理

3.1　渐近平直时空

自本节开始，我们将介绍一个新的专题——正质量定理（有时也称为正能量定理）。这一定理是经典广义相对论中一个很漂亮的结果，但在广义相对论的教材甚至专著中都极少介绍，读者要了解这一定理，往往只能求助于原始文献。而原始文献与教材或专著的一个很大的区别，就在于它的论述往往不是自给自足（self-contained）的，而要依赖其他原始文献。更麻烦的是，那些被依赖的原始文献本身——由于也是原始文献——又分别有自己所依赖的原始文献……这种层层依赖的结果是，除非读者已有足够的背景知识，否则为读懂一篇原始文献，往往要顺藤摸瓜地读上一大堆其他文献，其过程有如一个初学英语的人试图通过英-英词典查找词义，往往在词义之中又发现生词。对于正质量定理来说，这种困难在舍恩与丘成桐的论文中体现得尤为明显。在他们高度数学化的论文中，对物理背景及某些源自物理的数学表达式的由来交待得极为简略，或基本不做交待。不熟悉背景的读者哪怕细细研读他们的论文，也有可能只见树木，不见森林。因此在进入正题之前，我们将首先对正质量定理及其证明将会涉及的若干物理概念进行介绍。

引力场的能量动量问题一直是广义相对论研究中一个很困难的课题。自爱因斯坦以来，许多早期的广义相对论研究者，比如美国物理学家托曼（Richard Tolman，1881—1948）、苏联物理学家朗道、丹麦物理学家莫勒（Christian Møller，1904—1980）等，都曾在这一问题上做过很大努力。那些早期研究的目的之一，是想探究引力场本身的能量动量是如何分布的。现在回顾起来，那些研究虽然绝非毫无启示，但在很

大程度上归于了失败。后来的物理学家们大都认为，引力场的能量动量是不可定域的。从等效原理的角度讲，这一点几乎是显而易见的，因为对应于牛顿引力场的黎曼联络（Riemannian connection）可以通过坐标变换局域地消去。因此如果执意要寻找引力场能量动量分布的定域表述，就得付出很大的代价，比如限制坐标变换，引进高阶导数，放弃能量动量的协变性，设定背景度规或背景联络，等等。这些做法没有一种是省油的灯，而且过去多年的研究经验表明，即便付出那样的代价，结果依然不尽人意。美国物理学家米斯纳（Charles Misner，1932— ）、索恩（Kip Thorne，1940— ）和惠勒在其巨著《引力》（Gravitation）中曾经表示：寻找定域引力场能量动量的努力是“试图为一个错误的问题寻找正确的答案”。〔1〕

但另一方面，一对脉冲双星会因引力辐射而损失能量，从而导致轨道蜕变及轨道周期的变化，这种周期变化可以精确地加以计算，并获得观测的验证〔2〕。由于引力辐射是由纯引力场组成的，因此引力场本身携带能量动量是毫无疑问的事情。这样看来，问题的关键在于有关引力场的能量动量我们究竟可以知道多少，这是一个目前还在研究之中的问题。物理学界比较公认的一点是，一个孤立体系的总能量动量是可以定义的，这个总能量动量既包含了普通物质的贡献，也包含了引力场的贡献。对这一能量动量的理论表述是由美国物理学家阿诺维特（Richard Arnowitt，1928— ）、戴舍（Stanley Deser，1931— ）和米斯纳于20世纪60年代初提出的，被称为**ADM**能量动量，其中的能量部分则称为ADM能量（ADM energy），也叫**ADM**质量（ADM mass）。我们所要介绍的正质量定理涉及的就是ADM质量〔3〕。

要想定义一个孤立体系的总能量动量，首先必须搞明白什么是孤立体系。从物理上讲，一个体系的“孤立”指的是远离任何其他体系，或者更确切地说，是任何其他体系对它的影响都可以忽略。对于这样的体系，我们可以在各种距离上考察其物理性质，而不必担心受到其他体系

的影响。由于这一特点，我们可以利用该体系的某些与距离有关的物理性质来定义其“孤立”性。比方说，我们可以这样来定义一个孤立的带电体系：

（1）所有的源（电荷、电流）都分布在一个紧致的空间区域中；

（2）对于固定时刻t，沿空间方向远离体系——即空间距离$r\to\infty$——时场强的衰减与空间距离的平方成反比，即$F_{\mu\nu}\sim O(1/r^2)$；

（3）沿类光测地线远离体系时场强的衰减与空间距离成反比，即$F_{\mu\nu}\sim O(1/r)$。（请读者想一想，这一渐近行为的物理意义是什么？）

将这种思路运用到引力场中，一个很自然的设想是通过度规场$g_{\mu\nu}$以适当方式趋于闵科夫斯基度规$\eta_{\mu\nu}$来定义孤立体系。因为从物理上讲，远离一个孤立体系时，时空应该是平直的，这样的时空被称为渐近平直时空（asymptotically flat spacetime）。这种定义实质上是想通过时空的渐近平直性来定义孤立体系，这是现代广义相对论研究所采用的方法。在早期研究中，人们通常将渐近平直时空定义为度规满足条件

$$\begin{cases} g_{\mu\nu} \sim \eta_{\mu\nu} + O\left(\frac{1}{r}\right) \\ \partial_i g_{\mu\nu} \sim O\left(\frac{1}{r^2}\right) \end{cases} \qquad (3.1.1)$$

的时空（视情形不同，有时还会附加对$g_{\mu\nu}$更高阶导数的限制）。我们将会看到，这一定义所要求的渐近行为并非随意选取，而是出于确定引力场总能量动量的需要。不过，这一定义——我们称之为朴素定义——虽然在不少具体场合是适用的，对于普遍研究来说却有很大的局限性。这首先是因为它明显依赖于坐标的选择，比如定义中所用的空间距离r就是一个依赖于坐标选择的概念。在特殊的坐标下，$r\to\infty$甚至有可能只是一个有限远的点。不仅如此，用度规场在$r\to\infty$的行为来定义渐近平直

时空还有一个技术上的不利之处，那就是处理$r\to\infty$的极限并不是一件容易的事情，尤其是在不得不涉及多个极限或微分的相互次序的时候。

为了克服上述困难，自20世纪60年代起，彭罗斯等人提出了一个非常聪明的想法，那就是干脆把$r\to\infty$（即“无穷远”）这个“麻烦制造者”当作边界加入到时空流形中。这样一来，度规的渐近行为就可以用其在流形边界及其邻域内的微分性质来取代，从而避免采用像$r\to\infty$这样依赖于坐标选择的极限。当然，这个想法说起来容易，具体做起来却有不少微妙的细节需要处理。彭罗斯最初的工作只考虑了对类光无穷远的处理，在他之后，经过杰罗奇、美国加州大学圣塔芭芭拉分校（University of California，Santa Barbara）的物理学家霍罗威茨（Gary Horowitz）、印度物理学家阿什提卡（Abhay Ashtekar，1949— ）等多位物理学家的努力，直到20世纪七八十年代，人们才得到了渐近平直时空的相对完整的现代定义。

虽然我们所要介绍的正质量定理的表述和证明只需用到渐近平直时空的朴素定义，但我们仍将对现代定义的思路做一个简单介绍。这不仅是因为这一思路本身值得了解，而且也是因为通过对现代定义的介绍，我们可以对朴素定义的适用性有一个了解。如上所述，现代定义的关键是将“无穷远”纳入时空边界，要做到这一点，第一步显然是要对时空进行“压缩”，以便把看不见摸不着的“无穷远”拉到看得见摸得着的“有限远”。而要想对时空进行“压缩”，就必须改变时空的尺度。在数学上，这是通过所谓的共形变换（conformal transformation）$\bar{g}_{\mu\nu}=\Omega^2 g_{\mu\nu}$来实现的。这种变换之所以被称为共形变换，是因为（请读者自行证明）它只改变尺度而不改变角度，从而也不改变几何形状，其中包括光锥的形状。在相对论中，时空的因果结构是由光锥决定的，因此共形变换不改变时空的因果结构，这是它在广义相对论研究中广受青睐的根本原因。经过适当的共形变换（请读者想一想，为了将“无穷远”拉到“有限

远”，共形因子Ω需要满足的最基本的条件是什么？），再辅以一定的坐标变换，时空流形可以最终用一组在有限区间内取值的坐标来描述，这样就完成了将“无穷远”拉到“有限远”的任务。

完成了这一步之后，我们就可以为时空流形添加边界，那些边界点表示的就是原先可望不可及的“无穷远”，在这里被称为共形无穷远（conformal infinity），它们可分为：过去（未来）类时无穷i^-（i^+），过去（未来）类光无穷j^-（j^+），以及类空无穷i^0。在时空流形上添加边界所得到的结果有时被称为“非物理时空”（unphysical spacetime）[4]，得到这一“非物理时空”的过程则被称为共形紧致化（conformal compactification）。熟悉复变函数论的读者可能注意到了，对时空流形的这种处理方式类似于复变函数论中引进黎曼球面及无穷远点的做法。

为了更好地理解共形紧致化，我们来看一个简单的例子。我们知道，闵科夫斯基度规（在有关正质量定理的这一专题中，我们将采取一个与本书其余部分相反的约定，即将度规的空间部分取为正定，以便更方便地表示空间距离及类空超曲面的性质）

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega^2 \qquad (3.1.2)$$

所涉及的坐标r和t的取值范围都是无界的。但如果我们对这一度规接连实施下列三个变换：

（1）坐标变换$u=t-r$，$v=t+r$

（2）共形变换$\bar{g}_{\mu\nu}=[4/(1+u^2)(1+v^2)]g_{\mu\nu}$

（3）坐标变换$T=\arctan v+\arctan u$，$R=\arctan v-\arctan u$

就可将之转变为（请读者自行验证）

$$ds^2 = -dT^2 + dR^2 + \sin^2 R d\Omega^2 \qquad (3.1.3)$$

其中R和T的取值范围满足：$-\pi<T-R<\pi$，$-\pi<T+R<\pi$，$R\geq0$。这时所有坐标的取值全都成为了有界，这就完成了前面介绍的将“无穷远”拉

到“有限远”的任务。这一度规所对应的共形无穷远（即代表“无穷远”的边界）则是：过去类时无穷远i^-为$R=0$，$T=-\pi$，未来类时无穷远i^+为$R=0$，$T=\pi$，两者均为一个点；过去类光无穷远j^-为$T=R-\pi$，$0<R<\pi$，未来类光无穷远j^+为$T=\pi-R$，$0<R<\pi$，两者均为三维表面（拓扑结构为$S^2\times R^1$）；类空无穷远i^0为$T=0$，$R=\pi$，是一个点。

闵科夫斯基度规的这个例子虽然简单，但从中我们可以看到普遍定义所需处理的一个微妙的细节，那就是类空无穷远被映射为了一个点i^0。我们知道，物理量（尤其是带分量的物理量，比如矢量或张量）沿不同方向趋于类空无穷远所具有的极限往往是不同的（请读者举出一个具体例子），要想让这些不同的极限与单一的类空无穷远点相对应，“非物理时空”在这一点上的解析性质必须足够弱〔5〕。此外，在有物质存在的情形下，由于物质不会凭空产生和消失，因此我们不能要求时空在类时无穷远i^-和i^+趋于本质上是描述真空的闵科夫斯基时空。除这些细节外，渐近平直时空的定义还需满足以下三个物理上显而易见的条件：

（1）适当的完备性条件，以确保所有的无穷远都被包含在内。

（2）物质只分布在有限的空间区域内。这一条件原则上可以放宽为对物质能量动量张量$T_{\mu\nu}$在边界附近趋于零的方式的限制。

（3）度规张量在类空无穷远的某个邻域内趋于闵科夫斯基度规。这一条件所体现的是渐近平直的核心含义，它可以被巧妙地表示为Ω和$\bar{g}_{\mu\nu}$在边界上的约束条件，这一点也正是现代定义的优势所在。

渐近平直时空的现代定义就是由上述所有条件的数学表述组成的。但光有这样的表述还不够，我们还必须证明这一定义是无歧义的。这一点之所以需要证明，是因为满足定义要求的共形因子Ω是不唯一的。由于Ω的选择直接影响到时空边界（即“无穷远”）的结构，因此Ω的不唯一性原则上有可能导致所定义的时空边界（即“无穷远”）出现歧义。幸

运的是，杰罗奇等人早已证明，不同的Ω（当然它们首先得满足定义所要求的各种解析性质）所对应的“非物理时空”之间必定可以找到在物理时空上为恒等映射的微分同胚。存在这样的微分同胚表明，Ω的选择不影响时空的拓扑及微分性质，从而不会导致歧义〔6〕。这样就确立了渐近平直时空的现代定义。

渐近平直时空——以及以之为基础的孤立体系——的现代定义虽然抽象，并且看上去与早期研究所用的朴素定义很不相同，但上述条件中的第三条意味着，我们总可以找到一组适当的坐标，使度规张量在“远处”（即类空无穷远的某个邻域内）以朴素定义所要求的方式趋于闵科夫斯基度规。这表明，渐近平直时空的朴素定义虽不严格，但在实际应用中的确是能够行之有效的。

3.2　广义相对论的动力学

有了孤立体系的定义，下面我们来讨论孤立体系的总能量动量。由于能量动量是描述体系动力学行为的物理量，因此在定义它之前，有必要对广义相对论的动力学有所了解。我们知道，爱因斯坦场方程的左边包含时空度规$g_{\mu\nu}$及其一、二阶导数，右边则是描述物质分布的能量动量张量$T_{\mu\nu}$。场方程的这种形式导致了一种流传很广的误解，即以为所谓广义相对论的动力学，就是在时空流形上给定$T_{\mu\nu}$，求解$g_{\mu\nu}$。这种说法即便不算完全错误，起码也是似是而非的。因为时空流形是四维的，因此所谓“在时空流形上给定$T_{\mu\nu}$”，实际上包括了给出$T_{\mu\nu}$作为时间的函数。但所谓动力学，它的目的就是寻找物理量——无论其描述的是几何还是物质——随时间的演化，既然如此，又怎能事先就给定$T_{\mu\nu}$作为时间的函数呢？这既没有现实可行性，也不符合动力学的要求。

那么广义相对论的动力学究竟该如何定义呢？我们可以回想一下普通力学。在普通力学中，要解决一个动力学问题，往往需要给定所谓的初始条件——即一组动力学变量及其时间导数在初始时刻的空间分布。对于广义相对论来说，要想给出初始条件，首先要对时空进行某种分解，这样才能谈论所谓的“初始时刻”和“空间分布”。这种分解的实质，是将时空流形分解为$\Sigma^3\times R^1$，其中Σ^3（下文将省略维数上标）是三维类空超曲面，坐标记为x_i（i＝1，2，3，下同），表示空间；R^1则表示时间，坐标记为t。这种分解通常被称为3＋1分解，或**ADM**分解（Arnowitt-Deser-Misner decomposition）。在这样的分解下，每个时刻t都有一个对应的类空超曲面Σ。所谓初始条件，就是在某个给定时刻t_0（即所谓“初始时刻”），给出一组动力学变量及其时间导数在与该时刻相对应的类空超曲面Σ上的数值（即所谓“空间分布”）。

有了时空流形的分解，接下来就可以定义动力学变量。为此，我们

引进Σ的单位法矢量n，以及度规$g_{\mu\nu}$在Σ上的诱导度规$h_{ij}=g_{ij}+n_i n_j$。细心的读者可能注意到了，诱导度规h_{ij}其实就是2.2节中引进的时空度规的“空间部分”（不过两者之间有一个符号差异，请读者想一想，这个差异从何而来？）。h_{ij}作为Σ上的诱导度规，显然是动力学变量，h_{ij}在Σ上的分布则是初始条件的一部分。但光有h_{ij}的分布还不够，爱因斯坦场方程和大多数其他动力学方程一样，是二阶微分方程，因此我们还需要知道h_{ij}对时间的导数在Σ上的分布。不过由于ADM分解中的时间轴通常不与Σ正交，使用起来并不方便，因此人们往往用h_{ij}沿Σ法向n的李导数（Lie derivative），即

$$K_{ij}=\frac{1}{2}L_n h_{ij} \qquad (3.2.1)$$

来取代时间导数。这样定义的K_{ij}有什么好处呢？首先，它具有良好的坐标变换性质，这是由李导数的性质保证的；其次，它包含了h_{ij}的时间演化信息，从而的确可以取代h_{ij}对时间的导数来作为初始条件的一部分。事实上，时间基矢∂_t可以在n及（Σ上的）空间基矢∂_i组成的坐标系中分解为$\partial_t=Nn+N^i\partial_i$（其中N和$N^i$为分解系数），而$K_{ij}$则可以表示为

$$K_{ij}=\frac{1}{2N}(\partial_t h_{ij}-N_{i|j}-N_{j|i}) \qquad (3.2.2)$$

其中|i和|j表示相对于诱导度规h_{ij}的协变导数（下同）。从这一结果可以清楚地看到，K_{ij}包含了h_{ij}的时间演化信息（如果时间轴恰好沿n方向，那么K_{ij}将正比于h_{ij}的时间导数）。最后，但并非最不重要的，是K_{ij}具有清晰的几何意义。事实上，它是超曲面Σ的外曲率（extrinsic curvature），也叫做第二基本形式（second fundamental form），它描述的是Σ在它所嵌入的外部时空流形中的弯曲方式[7]。

为了更好地显示K_{ij}作为外曲率的几何意义，同时也为后文介绍舍恩与丘成桐的证明埋个伏笔，我们对K_{ij}再稍做一点考察。我们注意到，K_{ij}＝（1/2）$L_n h_{ij}$＝（1/2）$L_n g_{ij}=n_{(i;j)}$，其中$n_{(i;j)}$＝（1/2）（$n_{i;j}$

$+n_{j;i}$）是$n_{i;j}$的对称部分。不难看到，经过这样改写的K_{ij}与2.2节中测地线束形变的对称部分完全类似。事实上，我们这里所介绍的内容与那里有关测地线束的描述之间存在明显的对应关系：测地线的切矢V对应于n，与V垂直的子空间对应于Σ，形变对应于$n_{i;j}$，形变的对称部分则对应于K_{ij}。在2.3节中，我们提到过这样一个结果，即形变的反对称部分（即涡旋张量）在测地线束为超曲面垂直时为零。由于我们这里讨论的单位法矢量n和超曲面Σ显然满足超曲面垂直关系，因此形变$n_{i;j}$的反对称部分$n_{[i;j]}$为零。既然$n_{i;j}$的反对称部分为零，它就完全由对称部分所组成，而后者正是K_{ij}。这就表明

$$K_{ij} = n_{i;j} \tag{3.2.3}$$

这个结果的几何意义非常明确，它表明一个曲面（或超曲面）的外曲率所描述的是单位法矢量沿曲面运动时的变化，这完全符合我们从曲面所嵌入的外部空间（或时空）来判断其弯曲程度的直觉。

现在我们回到广义相对论的动力学上来。从上面的分析可以看到，广义相对论时空部分的动力学变量可以选为诱导度规h_{ij}（t，x^i），相应的初始条件则是初始时刻所对应的类空超曲面Σ上的诱导度规h_{ij}（x^i）及外曲率K_{ij}（x^i）[8]。从动力学变量与方程的数目对比来看，h_{ij}作为动力学变量是合理的，因为h_{ij}有6个独立分量，而爱因斯坦场方程虽然有10个方程，其中却只有6个是动力学方程，其余4个，即

$$G_{\mu\nu}n^{\nu} = 8\pi T_{\mu\nu}n^{\nu} \tag{3.2.4}$$

不包含度规对时间的二阶导数，因而只是对初始数据的约束条件。这些约束条件可以通过微分几何中所谓的高斯-柯达西方程组（Gauss-Codacci equations）表述为有关h_{ij}和K_{ij}的条件：

$$\begin{cases} {}^{(3)}R + (\mathrm{tr}K)^2 - \mathrm{tr}(K^2) = 16\pi\rho \\ [K_{ij} - (\mathrm{tr}K)h_{ij}]^{|j} = 8\pi J_i \end{cases} \tag{3.2.5}$$

其中[3] R是超曲面Σ上的曲率标量；tr是Σ上的迹；$\rho=T_{\mu\nu}n^{\mu}n^{\nu}$和$J_i=h_i^{\mu}T_{\mu\nu}n^{\nu}$则是物质能流密度的时间和空间分量。这是两个很重要的关系式，在舍恩与丘成桐的证明中将会被用到。

至此，我们就完成了对广义相对论的动力学变量及初始条件的介绍。所谓广义相对论动力学中的初值问题，指的就是给定一个类空超曲面Σ及其上满足约束条件（3.2.5）的h_{ij}、K_{ij}及物质分布，求解h_{ij}和K_{ij}在整个时空流形上的分布。对于我们所讨论的孤立体系来说，我们还要求Σ是渐近平直的类空超曲面。要做到这一点，Σ必须经过类空无穷远点i^0，并满足一定的解析条件[9]。在早期的工作中，类空超曲面Σ的渐近平直性与时空本身的渐近平直性一样，是通过朴素方式——即对h_{ij}、K_{ij}及其若干阶导数的渐近行为加以界定——来定义的。除此之外，人们往往还在时空流形上附加一定的因果条件，比如全局双曲条件（其定义可参阅2.4节），这些我们就不讨论了。

3.3　ADM能量动量

接下来我们介绍广义相对论的哈密顿表述（Hamiltonian formulation），ADM能量动量的原始定义就源于这一表述。我们知道，引力场的作用量密度是L＝（1/16π）（-g）$^{1/2}$R，其中g＝det（$g_{\mu\nu}$）是度规张量的行列式。由于（请读者自行证明）（-g）$^{1/2}$＝$Nh^{1/2}$，其中h＝det（h_{ij}）是诱导度规的行列式，以及（可以由高斯-柯达西方程组得到）R＝$^{(3)}$R＋tr（K^2）－（trK）^{2}，上述作用量密度可以表述为h_{ij}和K_{ij}的函数：

$$L = \frac{1}{16\pi} N h^{1/2} \left[{}^{(3)}R + \mathrm{tr}(K^2) - (\mathrm{tr}K)^2 \right] \qquad (3.3.1)$$

由此可知与h_{ij}对应的广义动量为

$$\pi^{ij} = \frac{\delta L}{\delta h_{ij}} = \frac{1}{16\pi} h^{1/2} \left[K^{ij} - (\mathrm{tr}K) h^{ij} \right] \qquad (3.3.2)$$

而引力场的作用量密度则可以改写为（其中丢弃了一些全微分项）

$$L = \pi^{ij} \partial_t h_{ij} - N R^0 - N_i R^i \qquad (3.3.3)$$

其中

$$\begin{cases} R^0 = -\dfrac{1}{16\pi} h^{1/2} \left[{}^{(3)}R - \dfrac{\mathrm{tr}(K^2)}{h} + \dfrac{(\mathrm{tr}K)^2}{2h} \right] \\ R^i = -\dfrac{1}{8\pi} \pi^{ij}{}_{|j} \end{cases} \qquad (3.3.4)$$

从这里我们可以看到，引力理论是一个有约束的理论，约束条件为R^0＝0和R^i＝0，N和N^i是相应的拉格朗日乘子。

由上述作用量密度可得引力场的哈密顿密度为

$$H = \pi^{ij}\partial_t h_{ij} - L = NR^0 + N_i R^i \qquad (3.3.5)$$

相应的哈密顿量（Hamiltonian）则为∫HdV（积分为类空超曲面Σ上的体积分）。至此，似乎一切都很顺利。但不幸的是，这个哈密顿量——作为h_{ij}和K_{ij}的泛函——可以描述引力场的动力学，却无法用来计算总能量，因为约束条件$R^0=0$和$R^i=0$的存在使它的数值恒等于零。要解决这个问题，必须注意到引力理论的一个微妙的细节，那就是在哈密顿量的变分中，有一些边界项在Σ渐近平直的情况下不为零。比如假定Σ的法矢量n在类空无穷远处渐近于时间平移生成元∂/∂t，即$N\to 1$，$N^i\to 0$，那么可以证明，哈密顿量的变分中有这样一个非零的边界项（它也是唯一一个非零的边界项）：

$$-\int \frac{1}{16\pi} N h^{1/2} h^{ij} \delta^{(3)} R_{ij} \mathrm{d}V \to -\delta \int \frac{1}{16\pi} (\partial^i h_{ij} - \partial_j h^i_i) \mathrm{d}S^j \qquad (3.3.6)$$

其中$\mathrm{d}S^j$是边界∂Σ上的面积元，积分理解为∂Σ趋于i^0的极限（下同）。显然，如果我们在哈密顿量上添加一个相应的边界项（1/16π）∫（$\partial^i h_{ij}-\partial_j h^i_i$）$\mathrm{d}S^j$，就可以抵消变分中的边界项。这一附加的边界项对动力学方程没有影响，却对总能量有贡献。事实上，由于哈密顿密度的体积分为零，这一边界项是对总能量的唯一贡献。这样，我们就得到了孤立体系的总能量〔10〕

$$E = \frac{1}{16\pi}\int (\partial^i h_{ij} - \partial_j h^i_i) \mathrm{d}S^j \qquad (3.3.7)$$

类似地，通过对N及N^i的渐近行为做其他假设，可以得到孤立体系总动量的表述：

$$p^i = \frac{1}{8\pi}\int (K^i_j - (\mathrm{tr}K)\delta^i_j) \mathrm{d}S^j \qquad (3.3.8)$$

可以证明，这样定义的E和p^i与Σ上渐近坐标的选择无关，并且构成类空

无穷远i^0上的四维矢量。这就是所谓孤立体系的**ADM**能量动量，其中的E被称为ADM能量或ADM质量，p^i被称为ADM动量。由上述定义不难看到，ADM能量动量的存在要求$\partial_k h_{ij} \sim O(1/r^2)$，而这正是渐近平直时空的朴素定义所要求的。

3.4　正质量定理

在上一节中，我们定义了ADM质量。在最简单的情况下，ADM质量是很容易计算的。作为练习，读者不妨计算一下施瓦西时空的ADM质量，以验证它与施瓦西度规中质量参数m的等同性。自ADM质量被提出以来，物理学家们一直有一个猜测，那就是ADM质量作为一个孤立体系的总能量，应该是非负的。这个猜测被称为正质量猜想。20世纪70年代末80年代初，正质量猜想被舍恩（曾经是丘成桐的学生）、丘成桐及美国数学物理学家威顿（Edward Witten，1951—　）所证明，从而成为了正质量定理。

正质量定理：一个孤立体系若其物质分布满足主能量条件，则其ADM质量非负。

在正质量定理中，孤立体系如我们在3.1节中所述，是通过渐近平直时空来定义的。为了使ADM质量本身有定义，正质量定理隐含了一个条件，那就是时空中存在具有良好解析性质的渐近平直三维类空超曲面Σ。正质量定理是一个有很大普适性的定理，这普适性主要体现在两个方面：一是物质分布只需满足主能量条件，而无须附加更精细的限制；二是Σ的拓扑结构可以有很大的复杂性，比如它可以包含不止一个渐近平直区域，并且各个渐近平直区域中的ADM质量可以彼此不同（这种时空的一个例子是我们以后将要介绍的虫洞）。另外要说明的是，所谓ADM质量非负，更“协变”地讲是指ADM能量动量非类空。

在文献中，正质量定理还包含另外一层含义（我们将不展开讨论），那就是如果ADM质量为零（在存在不止一个渐近平直区域的情形下，则如果ADM质量在其中一个渐近平直区域为零），那么Σ必定是三维欧几里德空间，而时空本身则必定是闵科夫斯基时空。这表明，闵

科夫斯基时空不仅是所有时空中能量最低的，而且是具有最低能量值的唯一时空。这一点对于确保闵科夫斯基时空的稳定性有很重要的意义〔11〕，同时它也意味着所有非平凡渐近平直时空的ADM质量都是正的，这是正质量定理中“正”字的含义所在。

正质量定理的证明花费了数学家和物理学家们整整20年的时间，它的证明者丘成桐和威顿由于其在证明这一定理及若干其他领域中的卓越贡献，于1982年和1990年先后获得了有“数学诺贝尔奖”之称的菲尔茨奖（Fields Medal）。这些事实都说明，正质量定理是一个艰深的数学问题，它的证明是第一流的数学成就。但另一方面，从物理上讲，一个孤立体系——尤其是经典意义下的孤立体系——的总能量非负难道不应该是显而易见的结果吗？事实上，在经典物理学的任何其他分支中，都从未出现过如此艰深的正质量定理，却唯独在广义相对论中出现了例外，这是为什么呢？定性地讲，一个孤立体系——哪怕是经典意义下的孤立体系——的总能量非负之所以不像想象的那样显而易见，是因为物质之间的引力结合能是负的。在牛顿引力理论中，一个孤立点质量产生的引力场的能量不仅是负的，而且可以趋于负无穷，这样的孤立体系的总能量也完全可以是负的〔12〕。由此可见，类似于正质量定理这样的结果在牛顿引力理论中是不存在的，孤立体系的总能量非负非但不是显而易见的结果，而且还是广义相对论有别于牛顿引力理论的一个重要特征（请读者想一想，电磁理论中也存在负的相互作用能，为什么却没有造成同样的困难？）。

细心的读者也许会提出这样一个问题：我们在本节一开始曾经提到，施瓦西时空的ADM质量正好等于施瓦西度规的质量参数m。如果我们将这个质量参数取为负的，即m＜0，不就可以得到一个负的ADM质量了吗？从单纯的数学计算上讲情况的确如此。但是，m＜0与m≥0的施瓦西时空之间存在着一些本质差异。对于我们所讨论的ADM质量来说，最重要的差异在于：在m＜0的施瓦西时空中，所有渐近平直三维

类空超曲面都会经过奇点r＝0，从而破坏ADM质量的定义（m≥0的施瓦西时空中则存在不经过奇点的渐近平直三维类空超曲面，从而使ADM质量有良好的定义）[13]。此外，我们在3.2节的末尾曾经提到，在研究广义相对论动力学的时候，人们往往在时空流形上附加一定的因果条件，比如全局双曲条件。m＜0的施瓦西时空很不幸地会破坏这类条件。从这个意义上讲，它的动力学本身就缺乏良好的定义，更遑论ADM质量。如果要追根溯源的话，那么这一切的困难都是由奇点r＝0造成的。对于m＜0的施瓦西时空来说，这一奇点是一种很特殊的奇点，被称为裸奇点，这是一种人见人怕（甚至连上帝都怕）的东西，我们将会在下一个专题——宇宙监督假设——中加以讨论。有读者可能会进一步问：既然问题出在奇点上，那么如果我们将点源换成非奇异的物质分布，是否就可以绕开这些问题呢？答案是否定的，因为m＜0的施瓦西时空无法与任何满足主能量条件的物质分布相匹配（从这里我们可以看到能量条件对于正质量定理的意义）。因此，m＜0的施瓦西时空并不满足正质量定理所要求的物理条件，从而不构成针对这一定理的反例。

就像数学或物理中几乎所有的困难问题一样，正质量定理在其最终证明出现之前，就曾经有数学家和物理学家给出过不完整或特殊情形下的证明。这方面最早的工作甚至出现在有关ADM质量的主要论文发表之前。早在1959年，美国普林斯顿大学（Princeton University）当时还刚刚成为博士的物理学家布里尔（Dieter Brill）就在Σ上的初始条件轴对称的情形下，对一类特定的引力波进行了研究（这类引力波有时被后人称为布里尔波），并证明了其能量非负。1968年，布里尔与戴舍（即“ADM”中的“D”）合作证明了当Σ上的初始条件与平坦条件（即闵科夫斯基时空下的初始条件）之差的二阶以上效应可以忽略时，正质量定理成立（这一结果正是我们在前面注释中提到的结果，即闵科夫斯基时空起码是准稳定的）。20世纪70年代，雷波维茨（Clement Leibovitz）、伊斯雷尔（Werner Israel，1931—　）、米斯纳

（即“ADM”中的“M”）等人先后证明了在Σ上的初始条件球对称的情形下，正质量定理成立。1976年，法国数学家肖盖（Yvonne Choquet-Bruhat，1923— ）和加拿大数学家马斯登（Jerrold Marsden，1942—2010）对布里尔和戴舍1968年的结果进行了推广，证明了当Σ上的初始条件与平坦条件之差在某种特定的泛函分析意义上“足够小”（该条件比布里尔和戴舍所要求的“二阶以上效应可以忽略”更弱）时，正质量定理成立。而1977年，莱特（Maria Leite）则给出了Σ可以等度规嵌入R^4这一特殊情况下正质量定理的证明。上面这些研究都是在纯广义相对论的范围内进行的，其中比较数学化的工作都侧重于从几何角度进行分析，在这一方向上的集大成者是舍恩与丘成桐，他们最终给出了正质量定理的第一种完整证明。

除上述结果外，戴舍和智利物理学家泰特伯姆（Claudio Teitelboim，1947— ）于1977年在当时还很新颖的超引力（supergravity）理论中发现了一个对于正质量定理的研究有着重要意义的结果，那就是超引力理论的ADM质量非负。这个结果对于广义相对论研究者来说很有些出乎意料，因为超引力理论从总体上讲是一个远比广义相对论复杂的理论，但就ADM质量非负这一特定结果而言，其证明从表观上看，却很奇妙地要比广义相对论中正质量定理的证明容易得多〔14〕。之所以出现这样的结果，是由于在超引力理论中，能量算符可以表示为在纯广义相对论中不存在的旋量性超荷算符（supercharge）的平方，从而具有形式上的非负性。1978年，格里沙鲁（Marc Grisaru）猜测这一结果可以在适当的极限下过渡为广义相对论中的正质量定理。超引力理论的这一特点——尤其是格里沙鲁的猜测——给了威顿很大的启示，使之于1981年给出了正质量定理的第二种证明。

3.5　舍恩与丘成桐的证明概述

从几何角度对正质量定理的研究自1979年起进入了一个大收获的时期。那一年，舍恩与丘成桐发表了一篇长达32页的论文，在假定Σ满足一个特定条件trK＝0的基础上证明了正质量定理。

trK＝0这一条件在正质量定理的研究中是很常用的。我们在上节中提到的很多工作，从布里尔在20世纪50年代末的工作，到布里尔和戴舍、肖盖和马斯登在20世纪60—70年代的工作，都用到了这一条件。我们首先来看一看，这个条件的几何意义是什么？

读者也许还记得，我们在3.2节中介绍K_{ij}的几何意义时，曾表示要“为后文介绍舍恩与丘成桐的证明埋个伏笔”。这个“伏笔”是什么呢？它就是我们在那里介绍的K_{ij}与测地线束形变之间的相似性，以及$K_{ij}＝n_{i;j}$这一结果。从这些结果立刻可以看到，trK——有时被称为平均外曲率（mean extrinsic curvature）[15]——作为$n_{i;j}$的迹，对应的是测地线束形变中描述测地线束会聚或发散趋势的膨胀标量θ（参阅2.2节），而trK＝0表示的显然是测地线束既不会聚也不发散，换句话说，其体积取极值。这个类比表明，Σ（处处）满足trK＝0这一条件意味着三维类空超曲面Σ的体积取极值。进一步的研究表明，这一极值是极小值。在微分几何中，人们将（处处）满足trK＝0的曲面称为极小曲面，但在广义相对论的研究中，它往往被称为极大超曲面（请读者想一想，为什么会出现这样的命名差异？）。因此，被舍恩、丘成桐以及其他研究者所广泛采用的这一条件可以表述为：时空中存在渐近平直的三维类空极大超曲面。

由于舍恩和丘成桐的证明是几何证明，因此，他们需要将主能量条件表述为几何条件。这可以利用我们在3.2节中介绍过的式（3.2.5），即

$$\begin{cases} {}^{(3)}R + (\mathrm{tr}K)^2 - \mathrm{tr}(K^2) = 16\pi\rho \\ [K_{ij} - (\mathrm{tr}K)h_{ij}]^{|j} = 8\pi J_i \end{cases} \tag{3.5.1}$$

来做到。这两个关系式是舍恩和丘成桐的论文的起点，也是本章前两节为读者理解舍恩和丘成桐的论文所做的主要铺垫。利用这两个关系式，主能量条件所要求的物质能流密度矢量非类空$\rho \geq |J^i J_i|^{1/2}$可以被表示为有关几何结构$h_{ij}$和$K_{ij}$的约束条件。由于主能量条件意味着物质能量密度$\rho \geq 0$，在舍恩和丘成桐所考虑的trK＝0的情形下，这显然只有在${}^{(3)}R \geq 0$的情况下才能满足（请读者自行证明这一点）。因此在trK＝0的情形下，主能量条件要求${}^{(3)}R \geq 0$，这个几何条件在正质量定理的几何研究中有着重要作用，在舍恩和丘成桐的证明中也被反复用到。

舍恩和丘成桐的证明采用的是反证法的思路，即通过假定ADM质量小于零来推出矛盾，其过程大致分为三步[16]：首先，他们证明了如果ADM质量小于零，那么在Σ中可以构造出一个特殊的二维极小曲面S，它在一个紧致集之外满足${}^{(3)}R \geq 0$。在这一步中，他们用到的是Σ渐近平直这一特点，以及${}^{(3)}R \geq 0$这一来自主能量条件的推论。由于S是极小曲面，因此S的面积泛函的二次变分必定非负，利用这一点，舍恩和丘成桐——作为第二步——证明了S的高斯曲率K在曲面上的积分∫KdS＞0。在这一步中，他们再次用到了${}^{(3)}R \geq 0$这一几何条件，以及第一步所得到的在S上的一个紧致集之外${}^{(3)}R \geq 0$这一构造性质。最后，为了推出矛盾，舍恩和丘成桐用两种不同的方法——其中只用到了Σ的渐近平直性以及S的构造性质——证明了一个与∫KdS＞0完全相反的结果，即∫KdS≤0。这一矛盾的出现表明ADM质量小于零这一假设与证明过程中所用的其他假设不相容。由于证明过程中所用的其他假设都是正质量定理本身的假设（比如Σ的渐近平直性）或其推论（比如${}^{(3)}R \geq 0$），因此这一矛盾的出现表明在正质量定理所假设的条件下，ADM质量必须非负。这样舍恩和丘成桐就完成了在trK＝0这一特定条件下正质量定理的

证明。

舍恩和丘成桐的这一证明无疑是一个很重要的结果，但由于它依赖于trK＝0，或者说Σ是极大超曲面这一额外条件，因此与正质量定理的完整证明还有距离。如我们在前面所述，trK＝0这一条件在正质量定理的早期研究中是经常用到的，在某些研究中人们几乎已将之视为正质量定理的前提之一。在舍恩和丘成桐的证明问世前不久，肖盖等人曾对这一条件进行了研究，确立了它在一系列合理情形下的有效性。但是，这充其量只能说明这一条件具有一定的合理性，却不足以确立其普遍有效性。要想得到一个真正普遍的正质量定理，这一条件必须去除。但这一条件一旦去除，从主能量条件中就无法得到（3）R≥0这一在正质量定理的几何研究中被反复用到的几何条件了（请读者自己证明这一点），从而上面所介绍的舍恩和丘成桐的证明，以及他们之前某些其他人的工作，就都不复成立了。因此这看似一步之差的细节，实际上是牵一发动全身，其修补难度是非常大的。不过舍恩和丘成桐在发表1979年的论文时，已经对整个证明有了全面构想。他们在发表那篇论文的同一年还发表了一篇摘要性的短文，概述了在更普遍（即不要求trK＝0）的情况下证明正质量定理的思路。在正质量定理被完全证明后，人们有时会回溯到这一年，即1979年，将之作为正质量定理被证明的年份。但事实上，舍恩和丘成桐的证明细节直到两年后的1981年才正式发表。在同一年的早些时候，他们还发表了另一篇短文，将正质量定理的表述由ADM质量非负推广为ADM能量动量非类空。

舍恩和丘成桐的完整证明发表后的第二年，曾在相关领域中做过研究的美国宾州大学（University of Pennsylvania）的微分方程及微分几何学家卡斯丹（Jerry　Kazdan）撰写了一篇介绍正质量定理的文章，对于舍恩和丘成桐1981年的证明，卡斯丹写下了一句简短的评语："任何充分的概述都将太长。"现在回过头来品味四分之一个世纪前的这句评语，我们感觉到情况并未发生太大变化，舍恩和丘成桐的证明在很大程

度上维持了当初的复杂性。因此我们将接受卡斯丹的忠告，不试图对它做任何“充分的概述”。舍恩和丘成桐在1981年的论文中所做的，用最简单——从而很不充分——的话来概述的话，是证明了在保持ADM质量不变的情况下，可以将普遍情形下的初始数据集（Σ，h_{ij}，K_{ij}）变形为满足$trK=0$这一条件的初始数据集。做到了这一点，普遍情形就被转化成了已经被证明的$trK=0$的情形。这样，舍恩和丘成桐就完成了正质量定理的证明。

3.6 威顿的证明概述

1981年对于正质量定理的证明来说是一个决胜之年。与舍恩和丘成桐发表完整的证明几乎同时，美国普林斯顿大学（Princeton University）的数学物理学家威顿也完成了一篇文章，给出了正质量定理的一种全新的证明方法。我们在3.4节中曾经提到过，威顿的证明受到了超引力理论中ADM质量非负这一结果的启发。超引力理论的结构相当复杂，它与广义相对论的最大差别，乃是除引力场外，还引进了作为引力子超对称伙伴（supersymmetric partner）的引力微子（gravitino）的场。由于引力子是一种玻色子（boson），而超对称理论中玻色子的超对称伙伴是费米子（fermion），因此引力微子是一种费米子，它的场是一种旋量场（spinorfield），它所满足的量子化条件则是反对易的，这意味着其“大小”正比于普朗克常数（Planck's constant）的平方根，从而在普朗克常数趋于零这一经典极限下应该能被忽略。有鉴于此，格里沙鲁——如3.4节所述——提出了一个猜测，那就是超引力理论的ADM质量非负这一结果可以在经典极限下过渡为一个经典广义相对论的结果。显然，这一结果——如果存在的话——应该就是正质量定理。格里沙鲁的这一猜测从未得到过证明，但它对威顿的证明起了很大的启发作用〔17〕。

威顿证明的出发点是Σ上的狄拉克方程（Dirac's equation）$\gamma^i D_i \psi=0$〔18〕，其中γ^i是狄拉克矩阵（Dirac matrices），D_i是四维协变导数在Σ上的分量，ψ是旋量场。威顿着重研究了这一方程的渐近于某个常数旋量ε——即$\psi \to \varepsilon+O(1/r)$的特殊解（这种特殊解有时被后人称为威顿旋量）。威顿之所以要研究这一狄拉克方程的这类特殊解，是因为这样的解出现在超引力理论中超荷算符的积分表达式中，这也正是超引力理论对威顿的启示所在。

利用Σ的渐近平直性，经过不太复杂的推理，威顿证明了对于任何ε≠0，上述狄拉克方程必定存在满足ψ→ε＋O（1/r）的特殊解ψ[19]。不过，威顿在证明这一存在性时用到了一些并非显而易见的结果，却未予严格论述。类似的不严密性在他的论文中还不止一处，有些甚至可以归为错误。这么多小缺陷同时出现在一篇论文中，对于数学功力极其深厚的威顿来说是颇为罕见的。因此从严密性上讲，威顿的原始证明与舍恩和丘成桐的证明有一定的差距。但幸运的是，威顿的证明发表之后，哈佛大学（Harvard University）的数学物理学家帕克（Thomas Parker）和陶布斯（Clifford Taubes，1954— ）很快就对他的证明做了改进，弥补了那些缺陷。

威顿的证明中另一个关键步骤，是将Σ上的狄拉克算符$\gamma^i D_i$与物质的能量动量分布联系起来。这一步的逻辑地位与舍恩和丘成桐的证明中那些将主能量条件表述为几何条件的关系式相类似，其重要性是显而易见的。因为只有建立了那样的联系，才能将主能量条件应用到证明中来。威顿在证明这一步时也出现了疏漏，忽略了狄拉克算符平方展开式中的一个曲率项，不过这一疏漏恰好被他在后文将狄拉克算符的平方展开式作用于ψ，并与ψ作内积时出现的另一个疏漏所抵消，因此未对整个论证造成实质性的破坏。而且更幸运的是，这一错误在文章付印前就被人发现，使威顿得以及时在文章中增添一个补注加以纠正。经过纠正后的这一联系可以表示为

$$(\gamma^i D_i)^2 = -D^i D_i + 4\pi G(T_{00} + T_{0j}\gamma^0\gamma^j) + K_i^j\gamma^0\gamma^i D_j \qquad (3.6.1)$$

其中最后一项中的K_i^j就是我们在3.2节中引进的Σ的外曲率（起初被威顿忽略的正是这一项）。这一结果其实是微分几何中的外森比克公式（Weitzenböck formula）应用于旋量场的情形。唯一的差别，是利用爱因斯坦场方程对某些几何量进行了替换，从而出现了与物质能量动量张量有关的项。

将上述结果作用于前面提到的特殊解ψ，显然可以得到

$$-D^iD_i\psi+4\pi G(T_{00}+T_{0j}\gamma^0\gamma^j)\psi+K_i^j\gamma^0\gamma^iD_j\psi=0 \quad (3.6.2)$$

用ψ与这一方程作Σ上的内积，并经过分部积分（正是在分部积分中，威顿出现了另一个疏漏，恰好抵消了他在狄拉克算符平方展开式中的疏漏），可以得到这样一个结果：

$$\int dS^k\psi^+D_k\psi=\int dVD^i\psi^+D_i\psi+4\pi G\int dV\psi^+(T_{00}+T_{0j}\gamma^0\gamma^j)\psi \quad (3.6.3)$$

在这一结果中，右端的第一项显然是非负的。而第二项由于$T_{0j}\gamma^0\gamma^j$的本征值为$\pm\|T_{0j}\|$，加上主能量条件要求$T_{00}\geq\|T_{0j}\|$，因此也是非负的。这表明

$$\int dS^k\psi^+D_k\psi\geqslant 0 \quad (3.6.4)$$

粗看起来，这个有关旋量场的不等式似乎与正质量定理没什么关系，但威顿注意到它左侧的边界积分是Σ上的不变量，而且除ψ的渐近值（常数旋量）ε外，它只与度规张量h_{ij}的O（1/r）渐近行为有关（更高阶的渐近项对面积分没有贡献）。另一方面，ADM能量E和动量p^i乃是由h_{ij}的O（1/r）渐近分量所能构成的仅有的不变量。因此威顿意识到这个不等式左侧的边界积分必定与ADM能量动量有关。那么它们之间究竟是什么样的关系呢？显然需要通过对这一面积分进行计算来揭示。在一般情况下，这将是一个很困难的计算，不过好在Σ具有渐近平直性，而面积分又处于渐近平直区域中，这使计算得到了极大的简化。计算的结果给出了一个相当简洁的关系式：

$$\int dS^k\psi^+D_k\psi=4\pi G\varepsilon^+(E+p_i\gamma^0\gamma^i)\varepsilon \quad (3.6.5)$$

因此前述不等式可以改写成

$$\varepsilon^+(E+p_i\gamma^0\gamma^i)\varepsilon\geqslant 0 \quad (3.6.6)$$

不难看到——与前面运用主能量条件证明$\int dS^k \psi^+ D_k \psi \geq 0$时所用的推理相类似——由于$p_i \gamma^0 \gamma^i$的本征值为$\pm\|p\|$，而且$\varepsilon \neq 0$，因此上述不等式意味着$E \geq \|p\|$，即ADM能量动量非类空。这样，威顿就完成了正质量定理的证明。正质量定理也因此有了两种截然不同的证明。

3.7 讨论

与数学物理中的很多其他定理一样，正质量定理的证明并不代表这一研究方向的终结。事实上，正质量定理被证明之后，很快就有物理学家将注意力转向了它的推广。在本节中我们将对这方面的工作做一个简单介绍。

威顿的证明发表两年后，英国理论物理学家吉本斯（Gary Gibbons，1946— ）等人就将他的证明推广到了渐近平直时空中包含黑洞的情形（舍恩与丘成桐的证明则无须推广就直接适用于吉本斯等人所考虑的情形）。他们并且还证明了，如果所涉及的黑洞含有电荷与磁荷，则ADM质量m≥$Q_e^2+Q_m^2$（其中Q_e和Q_m分别为总电荷与总磁荷）。吉本斯等人的证明完全沿用了威顿的方法，只是在协变导数中加入了电磁相互作用项。不过他们的结果有赖于一个附加条件，即电荷密度与磁荷密度的平方和不大于（不包括电磁场的）物质能量密度的平方。稍后，莫里斯奇（Osvaldo Moreschi）和斯帕林（George Sparling）将该结果进一步推广为一个带参数的不等式族。不过，他们的这一工作更多地只是一种纯粹的数学推广，而并无显著的物理意义。

渐近平直时空中包含有黑洞时，一个比吉本斯等人所考虑的更重要的命题是彭罗斯于1973年提出的彭罗斯猜想（Penrose Conjecture），也叫彭罗斯不等式（Penrose inequality），它可以表述为

$$M \geqslant \left(\frac{A}{16\pi}\right)^{1/2} \qquad (3.7.1)$$

其中M为ADM质量，A为所有黑洞的最外部视界面积之和。假如所讨论的黑洞为质量m的施瓦西黑洞，则A＝16πm^2，彭罗斯猜想可以简化为M≥m，即如果渐近平直时空中存在一个质量为m的施瓦西黑洞，则整个时空的ADM质量必定不小于m。与正质量定理所包含的ADM质量为零

意味着时空为闵科夫斯基时空这层含义相类似，彭罗斯猜想要求等号只在时空为施瓦西时空时才成立。显然，彭罗斯猜想比（包含黑洞的）正质量定理（M≥0）更强。在后文中我们将会看到，彭罗斯猜想与所谓的宇宙监督假设颇有渊源。

彭罗斯猜想提出二十几年后，2001年，数学物理学家们在证明一类特殊情形（即渐近平直超曲面的曲率标量R≥0的情形）下的彭罗斯猜想上取得了重要进展。我们在3.5节中介绍舍恩与丘成桐的证明时曾经介绍过R≥0这一条件，它是Σ为极大超曲面情形下主能量条件的推论。数学物理学家们通常把满足这一额外条件的彭罗斯猜想称为黎曼-彭罗斯猜想（Riemannian Penrose conjecture）〔20〕。1997年，惠斯肯（Gerhard Huisken）和伊尔玛能（Tom Ilmanen）在时空中只包含一个黑洞（确切地说是时空只包含一个单连通视界）的情形下证明了黎曼彭罗斯猜想。两年后（1999年），布雷（Hubert Bray）给出了更普遍情形下黎曼彭罗斯猜想的证明〔21〕。但迄今为止还没有人能够将R≥0这一额外条件去除，因此一般情形下的彭罗斯猜想仍未得到证明。

对正质量定理的另一类推广是试图将之推广到四维以上的时空。这也是一个迄今尚未完全解决的问题。在我们介绍过的两种正质量定理的证明中，舍恩和丘成桐的证明往高维方向可以直接推广到八维及八维以下时空，但对于高于八维的时空，他们的方法会遭遇迄今无人能够克服的奇点困扰〔22〕；而威顿的证明所依赖的旋量结构在高维时空中并不普遍存在，因而也不能直接推广到高维情形。尽管困难重重，人们仍在这方面作着艰难努力。2006年，洛卡姆（Joachim Lohkamp）发表了一篇文章，试图将舍恩和丘成桐的证明推广到八维以上时空。可惜的是，在这一工作中他不得不重新引进了R≥0这一阴魂不散的额外条件。这表明，即便他的证明本身无误，其结果也并不具有普遍性。

除上面提到的推广外，数学物理学家们还考虑了在某些特定的非渐近平直时空——比如渐近AdS时空（Anti-de Sitter spacetime）——中的

正质量定理。另外，除了舍恩和丘成桐以及威顿的经典证明外，也有人尝试用其他方法来证明正质量定理。比如彭罗斯、索金（Rafael Sorkin）和伍尔加（Eric Woolgar）在1993年提出过一种以时空的因果结构（具体地讲是以类光测地线的性质）为基础的新证明〔23〕。这些我们就不再介绍了。

注释

〔1〕不过，人们并未停止这方面的尝试，前面提到的那些可能的做法每一种都仍有人在研究。也有人在研究其他引力理论中引力场的能量动量分布。

〔2〕其中最著名的事例来自脉冲双星PSRB1913＋16，这也是人类最早发现的脉冲双星。1993年，它的发现者美国天体物理学家赫尔斯（Russell Hulse，1950—　）和泰勒（Joseph Taylor，1941—　）获得了诺贝尔物理学奖。

〔3〕除ADM质量外，在广义相对论研究中还会用到其他一些质量概念，比如邦迪（Bondi）质量。对于那些质量也存在相应的正质量问题或正质量定理。限于篇幅，本章将只讨论ADM质量。

〔4〕这里虽然用了“非物理”这一称谓，但事实上“非物理时空”的边界完全是由物理时空的共形性质所决定的，从这个意义上讲它还算“物理”。

〔5〕以度规张量为例，它在i^0只具有连续性以及一阶导数对方向的连续依赖性。

〔6〕需要提醒读者的是，虽然渐近平直时空的定义本身不存在歧义，但不同物理文献所采用的定义彼此间却往往是互有歧义的，即并不存在唯一的标准定义。不过那些定义之间的细微差异对于多数侧重物理的研究来说并无重大影响。

〔7〕了解外曲率的一个很好的例子是圆柱面，许多初学者在学习微分几何的时候，恐怕会对圆柱面的曲率张量为零感到困惑。因为在直觉上，圆柱面分明是“弯曲”的。这一直觉其实没有错，只不过它是来自于圆柱面所嵌入的三维空间，而非圆柱面本身，因而体现的是圆柱面的外曲率，而非内蕴曲率。倘若我们不是生活在三维空间，而是生活在圆柱面上，并且只能局域地感知圆柱面，那就不会觉察出它与平面的差异，从而也就不会产生诸如“圆柱面是弯曲的”那样的直觉。

〔8〕有读者或许会问：出现在K_{ij}中的时间基矢的分解系数N和N^i是否也是动力学变量？答案是否定的，因为我们在3.3节中即将看到，广义相对论的动力学是有约束的动力学，N和N^i是与约束条件相对应的拉格朗日（Lagrange）乘子。

〔9〕人们通常将三元组（Σ，h_{ij}，K_{ij}）称为初始数据集（initial data set）。而如果Σ渐近平直，且在i^0满足一定的解析条件（这种解析条件对物质场的渐近行为也是一种约束），则相应的初始数据集被称为渐近平直的初始数据集。

〔10〕这一能量（即ADM能量或ADM质量）虽然来自于对引力场作用量的分析，但它包含了物质场的贡献，因为后者已体现在了h_{ij}的渐近行为中（在弱场近似下，这一点可以通过直接计算予以证实）。ADM质量由渐近平直区域中的面积分表示，体现了一个朴素并且实用的物理思想，那就是确定一个孤立体系总能量的方法，是从远处观测其引力效应。另外，从这一表述可以看到，如果Σ是一个封闭曲面（比如S^3），那么由于其边界为零，E必定为零。人们有时把这一结果表述为：封闭宇宙的总能量为零。不过，这种表述的物理意义是值得怀疑的，因为在不存在渐近平直区域的情况下，总能量这一概念本身就很难有合适的定义。

〔11〕早在正质量定理被证明之前，人们就已经知道，闵科夫斯基时空起码是准稳定的（quasi-stable），即它的能量相对于其他时空是（泛函意义下的）局部极小值。这是因为对闵科夫斯基时空的二阶微扰是所谓的引力波，它的能量早已被证明是正的。不过，如果正质量定理不成立，那么闵科夫斯基时空仍有可能是不稳定的，因为它有可能会通过非微扰过程衰变为ADM质量为负的时空。

〔12〕这里我们假定牛顿引力只取决于物质，而不受引力场自身能量的影响，这是牛顿引力作为线性理论的重要特点。

〔13〕我们在2.1节中曾经说过，奇点并不存在于物理时空中，因此这里所说的“经过奇点”实际上是指无限接近（但不包含）奇点。感兴趣的读者可以思考一下，渐近平直三维类空超曲面无限接近（但不包含）奇点r＝0为什么会破坏ADM质量的定义？

〔14〕严格来讲，超引力理论中的正质量定理只是在形式上得到了证明，它的真正证明其实还根本不存在。这是因为超引力理论既不是一个经典理论，也不是一个可重整的量子理论，它的真正含义，以及它的包括能量动量在内的许多物理量的严格定义，直到今天也还是有待澄清的问题。因此只能说其证明“从表观上看”容易得多。

〔15〕平均外曲率的定义有时与trK差一个常系数。

〔16〕这里所述的步骤与舍恩和丘成桐的原始论文对证明步骤的分解略有差异，本文所述的第一步相当于舍恩和丘成桐原始论文中前两步的合并，而舍恩和丘成桐原始论文中的第三步则被我们分解成了两步。

〔17〕需要强调的是，超引力理论及格里沙鲁的猜测对威顿的证明虽然有着启示作用，但威顿的证明并不依赖于超引力理论或格里沙鲁的猜测在现实世界中的适用性。也正因为如此，尽管超引力理论的ADM质量非负——如我们在3.4节中所说——只是一个形式上的结果，它能在经典极限下过渡为经典广义相对论中的正质量定理也只是一个猜测，但威顿对正质量定理的证明却是一个严格证明。

〔18〕这里略去了对我们的介绍没有实质影响的常数因子，不同文献对这类因子的约定各不相同。

〔19〕如果ε＝0，则可以证明不存在满足上述狄拉克方程的非平凡解。

〔20〕相应地，满足这一额外条件的正质量定理也常被称为黎曼正质量定理。不过德国数学家黎曼（Bernhard Riemann，1826—1866）跟这些冠着他名字的结果并没有直接关系。

〔21〕这里所说的时间（即1997年和1999年）是这两项工作实际完成的时间。它们直到2001年才正式发表在知名学术刊物《微分几何杂志》（Journal of Differential Geometry）上，其中布雷的文章发表在9月刊上，惠斯肯和伊尔玛能的文章发表在10月刊上。

〔22〕2007年5月，布雷与李（Dan Lee）采用了类似于舍恩和丘成桐的技巧，将布雷本人所证明的黎曼彭罗斯猜想也推广到了八维及八维以下时空。

〔23〕他们的这一工作响应者寥寥无几，除作者之一的伍尔加本人曾试图对这一证明进行推广外，极少有其他文献提及这一工作。

第4章
宇宙监督假设

4.1 黑洞、裸奇点及宇宙监督假设

在2.5节中，我们介绍了霍金-彭罗斯奇点定理。按照这一定理，奇点的形成在经典广义相对论中几乎是无可避免的。但奇点的存在对广义相对论的动力学有着很消极的影响。以最常见的施瓦西奇点为例，时空的曲率在奇点附近会趋于发散，物理学定律在那样的极端条件下将不再适用。即便对于性情比较“温和”——即时空曲率不发散——的奇点，它所具有的测地不完备性可以导致粒子在有限时间内从时空流形中消失，对于传统的物理学定律来说依然是一种破坏。虽然奇点本身——如我们在2.1节中所说——并不存在于物理时空之中，但如果由它造成的物理定律的破坏可以对物理时空中的演化产生影响，那么这种“借尸还魂”般的影响就足以使得我们无法有效地预言物理时空中的演化。

显然，奇点有可能具有的这种不良品性对广义相对论是一种威胁。因为对广义相对论来说，预言物理时空中的演化是一项重要使命。而从原则上破坏这种预言能力，则无疑是对广义相对论的一种巨大的，甚至堪称是颠覆性的破坏，这不仅是物理学家们不希望看到的，而且——在某些物理学家看来——也是上帝他老人家不希望看到的。

那么，在奇点本身几乎无可避免的情况下，有什么办法能够避免奇点有可能带来的破坏作用呢？这便是我们要在本节及接下来的几节中关注的问题。为了对这一问题及可能的解决方式有一个初步认识，我们先来看一个大家熟悉的例子：施瓦西奇点。我们知道，对于施瓦西奇点来说，大自然以一种非常有效的手段掩盖住了奇点所具有的不良品性，那就是：施瓦西奇点总是被包裹在所谓的施瓦西视界之内，而施瓦西视界

之内的区域（被称为施瓦西黑洞）与外部时空在因果上是完全隔绝的。这表明由施瓦西奇点所造成的任何物理定律的破坏都被严严实实地掩盖在了黑洞（或视界）之中，而不会影响外部时空中的演化。从这个意义上讲，施瓦西奇点的存在是无害的。

但施瓦西奇点只是最简单的奇点，它被包裹在黑洞（或视界）之内这一性质能在多大程度上代表奇点的一般性质？特别是，当施瓦西奇点所具有的各种对称性不复存在的时候，奇点是否仍具有这种被黑洞（或视界）所包围的特性？这些却都是未知数。一个奇点倘若不被黑洞（或视界）所包围，它所造成的物理定律的破坏就有可能对物理时空中的演化产生影响，那样的奇点被称为裸奇点。为避免奇点的存在破坏时空中的演化性质，裸奇点的存在是必须被排除的。而排除裸奇点的一个重要途径，就是证明黑洞（或视界）的出现具有普遍性，而不依赖于任何特殊的对称性。

读者也许还记得，我们在1.1节中曾经介绍过有关奇点的出现是否依赖于对称性的争论。奇点定理的证明很漂亮地解决了那一争论。现在的问题几乎是当年那场争论的翻版，只不过当年所争论的是奇点的产生是否依赖于对称性，而我们现在所讨论的则是黑洞（或视界）的产生是否依赖于对称性。很多物理学家希望对后者也能找到一个像当年奇点定理那样的普遍答案，即凡能产生奇点的物理条件，也一定能产生包裹奇点的黑洞（或视界）。显然，这样的答案倘若存在，我们就不必担心奇点的出现会破坏时空中的演化性质了。这一良好愿望被称为宇宙监督假设，它是1969年由彭罗斯提出的。通俗地讲，宇宙监督假设要求所有的奇点都受到黑洞（或视界）的“监督”（即不存在裸奇点）。它的铁杆支持者霍金曾将这一假设幽默地表述为：上帝憎恶裸奇点（God abhors a naked singularity）。

可惜的是，与奇点的形成不同的是，对于黑洞（或视界）的形成，目前还没有普遍的结果。当然，这也不难理解，因为我们将会看到，黑

洞（或视界）的定义与时空的整个未来演化有着密切关系，因而不像奇点的定义那样局域，研究黑洞（或视界）的形成也因而要比研究奇点的形成困难得多。有些物理学家甚至不无悲观地认为，对那些性质的研究超出了人们迄今在微分几何及微分方程方面所具有的能力。这种悲观看法是否有道理，还有待观察，目前物理学家们所知道的是：描述黑洞的某些时空——比如施瓦西时空和克尔（Kerr）时空——在线性微扰下是稳定的。这是20世纪70—80年代由维希外希瓦拉（C. V. Vishveshwara）、普赖斯（Richard Price，1943— ）、凯（Bernard Kay）、沃尔德（Robert Wald，1947— ）和怀廷（Bernard Whiting）等人所证明的。这些结果意味着如果物质分布非常接近于形成施瓦西时空或克尔时空所需要的对称性，则它们的坍缩将会产生黑洞而非裸奇点。这一结果虽有很大的局限性，但对宇宙监督假设乃是很重要的早期支持。

另一方面，虽然裸奇点不受欢迎，但我们必须看到，在普遍意义下摒弃裸奇点是不现实的。事实上，按照大爆炸宇宙论，我们所生活的宇宙始于约138亿年前的一次大爆炸。从经典广义相对论的意义上讲，大爆炸正是一个裸奇点，我们不仅可以观测到它的残留效应，而且我们及我们周围的一切本身就是残留效应的一部分。不过大爆炸乃是既成事实，只要此后的时空演化不再产生新的裸奇点，广义相对论在实际意义上的预言能力就不会受到破坏。因此，在考虑宇宙监督假设时，像大爆炸这样的初始奇点需要被排除在外[1]。宇宙监督假设关心的乃是在正常的物质性质及初始条件下，通过诸如引力坍缩之类的演化过程能否产生出裸奇点。

如果裸奇点不存在，那么所有的奇点就都应该被黑洞（或视界）所包围。因此在进一步讨论宇宙监督假设之前，让我们先对黑洞（或视界）的含义作一点讨论。在广义相对论中，黑洞这一概念是在渐近平直时空——或者更确切地说，是在所谓强渐近可预测（strongly

asymptotically predictable）时空——中进行定义的[2]。什么叫做强渐近可预测时空呢？它是一种特殊的渐近平直时空[3]，在其中类光无穷远的因果过去——通常记为J^{-}（j^{+}），其中J^{-}表示因果过去，j^{+}表示类光无穷远——是全局双曲的[4]。由于“类光无穷远的因果过去”是由所有可以与类光无穷远建立因果联系的时空点所组成，而“全局双曲”——如我们在2.4节中所说——意味着时空中的演化可以通过适当的初始条件来预言。因此在强渐近可预测时空中只要给定适当的初始条件，我们就可以对所有能用光信号到达类光无穷远的部分——即J^{-}（j^{+}）——做出完善的预言。

那么什么是黑洞呢？黑洞是由所有无法与类光无穷远建立因果联系的时空点所组成的时空区域（这正是“连光也无法从黑洞中逃逸”这一通俗说法的物理表述）。如果用M表示整个时空流形，那么所谓黑洞就是指$M-J^{-}$（j^{+}）（感兴趣的读者可以思考这样一个问题：我们把黑洞定义为由无法与类光无穷远建立因果联系的时空点组成的时空区域，这是否意味着黑洞可以与黑洞外有限远的时空点建立因果联系？）。黑洞的边界则被称为视界，或者确切地说是事件视界（event horizon），以区别于其他一些视界概念，比如2.4节中提到的柯西视界。显然，黑洞与视界的上述定义与时空的整个未来演化有着密切关系，因为对类光无穷远及能与之建立因果联系的点集的确定，都有赖于时空未来演化的整体性质。

将黑洞的概念与强渐近可预测时空的定义联系起来，我们可以看到，在强渐近可预测时空中，除黑洞以外的所有区域都具有良好的因果性质（全局双曲）。这样的区域显然不可能包含奇点。换句话说，在强渐近可预测时空中如果存在奇点，则奇点必定存在于黑洞之中，这正是宇宙监督假设所预期的性质[5]。因此，利用强渐近可预测时空这一概念，宇宙监督假设——确切地说是所谓的弱宇宙监督假设（weak cosmic censorship hypothesis），以区别于后文将会提到的强宇宙监督假

设（strong cosmic censorship hypothesis）——可以粗略地表示为：

宇宙监督假设（粗略版）：在正常的物质性质及初始条件下，时空是强渐近可预测的。

不过，这一表述虽然对宇宙监督假设的结论部分做了比较精确的表述（即时空是强渐近可预测的），却没有对前提部分，即所谓“正常的物质性质及初始条件”做出明确界定，因此充其量只是一个“半拉子工程”。那么，为了使宇宙监督假设成立，究竟什么才是所需要的“正常的物质性质及初始条件”呢？不幸的是，这是一个极其困难的问题——事实上，这个问题和宇宙监督假设本身一样，是一个迄今尚未完全解决的问题。但是，这个极其困难的问题对于宇宙监督假设的研究又是至关重要的，从而无法回避——因为若不假定“正常的”初始条件，就无法将初始奇点及其他一些奇巧的情形排除在外；而若不假定“正常的”物质性质，则可以轻而易举地用纯几何的方式构造出有裸奇点的时空，然后通过逆用爱因斯坦场方程来定义相应的物质分布（这样定义的物质分布通常具有奇特的性质），从而构造出宇宙监督假设的反例。从这个意义上讲，宇宙监督假设是一个双重难题，它的表述本身就是一个难题，它的证明则是难上加难。

值得一提的是，在探索宇宙监督假设的征途上，物理学家们除发挥智慧外，也充分展示了他们独特的幽默感。1991年9月24日，霍金与同事普雷斯基尔（John Preskill，1953—　）及索恩以宇宙监督假设为对象打了一个赌（当然，霍金把赌注压在了自己强力支持的宇宙监督假设的成立上）。按照赌约，输家必须向赢家提供足以覆盖后者裸体的衣服（看来奇点是否受到监督虽还不得而知，赢家的裸体是必须受到衣服“监督”的），并在衣服上绣上愿赌服输的话。霍金这人赌运总体上讲是比较差的，但他的赌品倒是可圈可点。6年后，随着一些我们在后文中将会介绍的对宇宙监督假设不利的证据得到确立，霍金公开承认了自

己的落败。不过，在盘点了赌场失意的原因后霍金表示，自己之所以失利，乃是因为在赌约——即对宇宙监督假设的表述——中考虑不周，未对初始条件——即我们上文提到的“正常的初始条件”——作出足够的限定。换句话说，霍金认为自己是栽在了我们上文提到的“半拉子工程”上。当然，这番辩白在赢家普雷斯基尔及索恩看来纯属借口，于是这三位老顽童在霍金对赌约略作调整的基础上重开了赌局。这一年，即1997年，他们打赌的消息出现在了《纽约时报》（The New York Times）的科学专栏中，成为了大众新闻。

在接下来的几节中，我们将讨论一些与宇宙监督假设有关的论据，其中有些论据有利于宇宙监督假设，有些则不利。我们将通过对这些论据（尤其是不利论据）的讨论，来试图完善宇宙监督假设的表述。

4.2　摧毁黑洞——不可能的任务？

对宇宙监督假设的早期考察涉及一类很特殊的努力，那就是摧毁黑洞（或视界，下同）——当然只是通过理想实验。那种努力的基本思路是：既然宇宙监督假设要求所有奇点都被黑洞所包围，那么假如有办法在不破坏奇点的情况下将黑洞摧毁，原先遁迹其中的奇点岂不就暴露在了“光天化日”之下，从而变成裸奇点了？如果那样，宇宙监督假设就被推翻了。对于一个假设，能证明它成立固然很好，但推翻它显然也是一种解决方式。而且反过来说，即便无法推翻，只要努力“推”了，也算是积累了证据。因此，这种努力不失为考察宇宙监督假设的一个直观并且有益的研究方向。

那么，黑洞有可能被摧毁吗？

摧毁黑洞，这堪称是前所未闻的故事。事实上，在只有施瓦西解的情形下，这个问题或许连提都不会被提出。因为我们都知道，由施瓦西解所描述的施瓦西黑洞在经典广义相对论中是不可能被摧毁的。它就像一个贪婪而吝啬的守财奴，只认得一件东西，那就是质量。而且它对质量向来是只知索取，却绝不付出，它的视界当然也绝不会凭空消失，而只会像守财奴的钱包一样越来越大。但是，施瓦西黑洞只是最简单的黑洞，自施瓦西黑洞之后，物理学家们在广义相对论中又陆续发现了一些更复杂的黑洞解，它们不仅呈现出丰富多彩的性质，而且也为摧毁黑洞提供了一些可能的途径——当然，只是“可能”而已。

那么，那些更复杂的黑洞是什么样子的呢？研究表明，对于引力与电磁耦合的所谓爱因斯坦-麦克斯韦体系——也称为电真空（electrovac）体系，广义相对论的所有稳定解（stationary solution）都可以由所谓的克尔-纽曼（Kerr-Newman）度规所描述〔6〕。这一度规只带三个参数：质量m、电荷Q，以及单位质量所带的角动量J。这一结果

被称为黑洞无毛发定理（no hair theorem），或黑洞唯一性定理（uniqueness theorem）。对于我们的目的来说，克尔-纽曼度规的一个令人瞩目的特点是它有两个视界，分别位于径向坐标

$$r_{\pm} = m \pm (m^2 - Q^2 - J^2)^{1/2} \qquad (4.2.1)$$

处。其中位于$r=r_+$的被称为外视界，它是事件视界，这一视界以内的区域被称为克尔-纽曼黑洞；而位于$r=r_-$的被称为内视界，它是柯西视界。假如电荷与角动量都为零，则内视界消失，外视界等同于施瓦西视界。除此之外，克尔-纽曼度规还有一个重要特点，那就是在赤道面上有一个等效半径为J的奇环。

由上述视界半径公式，细心的读者也许自己就能看出一个问题，那就是内外视界在$Q^2+J^2=m^2$时将趋于重合，而在$Q^2+J^2>m^2$时则会失去意义，因为视界公式中的被开方式将会变成负数。从物理上讲，这时候克尔-纽曼度规将不存在视界，从而也将不再描述黑洞。但克尔-纽曼度规所具有的奇环却依然存在，这个奇环此时就变成了所谓的裸奇环，它的出现将破坏宇宙监督假设。这对于我们雄心勃勃的摧毁黑洞计划来说，无疑是一条重要线索。看来贪婪的黑洞似乎也有弱点，这弱点正是贪婪！通俗地讲，我们只要设法让一个它吃进过量的电荷或角动量，使得$Q^2+J^2>m^2$，就能把它“撑死”。

这计谋看起来不错，但关键是：有可能得逞吗？这种源源不断地向黑洞输送电荷或角动量，直至将之摧毁的过程真的有可能实现吗？关于这一点，物理学家们曾经作过分析，但结果——很遗憾地——却是否定的。否定的理由其实很直观，下面介绍一下。为简单起见，我们只考虑通过获得过量的电荷来摧毁一个不旋转黑洞——即所谓的雷斯勒-诺斯特朗姆（Reissner-Nordström）黑洞——的努力。假定我们每次向黑洞投放电荷δQ，携带这一电荷的粒子质量则为δm。为了让黑洞的电荷增加快于质量增加，我们要求$\delta Q>\delta m$。向黑洞投放这样的电荷在一开始是

很容易的，事实上，我们只要将电荷放在黑洞周围，它就会自动地被黑洞的引力所俘获。可惜这样的好光景并不能持久，随着黑洞的总电量Q越来越接近总质量m，继续向黑洞投放电荷就会变得越来越困难，因为黑洞中已有的电荷将会对新投放的电荷δQ产生越来越强烈的排斥，这种排斥最终将会超过质量δm所受的引力（请读者想一想为什么）。在这种情况下，为了让新投放的电荷能被黑洞俘获，我们就不能简单地将电荷放在黑洞周围，而必须使劲地将它扔向黑洞，让它依靠初速度来克服来自黑洞电荷的排斥作用。但是这样做的副作用却是增加了携带电荷的粒子所具有的初始能量，即增加了δm。1974年，沃尔德计算了这一初始能量的大小，结果发现若所讨论的黑洞已经处于Q＝m的极端带电状态，那么向这一黑洞投放电荷δQ所需提供的初始能量δm必须满足

$$\delta m \geqslant \delta Q \qquad (4.2.2)$$

由此可见，为了能在这种极端情形下继续向黑洞输送电荷，被输送的电荷所具有的能量将会自动保证m≥Q继续得到满足。换句话说，黑洞是不会因为上面这种输送电荷的方法而被摧毁的。

类似地，如果我们试图通过向一个J＝m的极端旋转黑洞投送具有很大轨道角动量的粒子以突破m≥J，则该粒子所具有的碰撞参数（impact parameter）将大到使之无法击中黑洞。而如果我们试图投送的是一个具有很大自转角动量的粒子，则该粒子与旋转黑洞之间会产生自旋-自旋相互作用，这种相互作用同样会阻止其击中黑洞。将上面这些分析综合起来，我们看到，通过向黑洞投放电荷或角动量来摧毁黑洞乃是不可能任务（mission impossible）[7]。大自然在极端黑洞情形下显示出的保护黑洞的手段是令人瞩目的，简直像是在有预谋地阻止裸奇点的出现，这是支持宇宙监督假设的一类很直接的证据[8]。

4.3　彭罗斯猜想与宇宙监督假设

对破坏宇宙监督假设的可能性的另一类考察是彭罗斯所做的。他考虑了一个渐近平直超曲面S上满足主能量条件的某种物质分布〔9〕。视具体情况而定，这种物质分布有可能产生封闭陷获面（封闭陷获面的定义请参阅2.4节）。如果宇宙监督假设成立，那么可以证明，所有这些封闭陷获面都必须被包含在黑洞之内，我们用A表示黑洞的外视界面积。在物理上可以预期，这一黑洞在经历了各种经典广义相对论所允许的变化——比如吸积物质，辐射引力波，等等——之后，最终将会进入某种稳定状态，即变成一个稳定的黑洞。按照黑洞无毛发定理，这一稳定的黑洞必定是克尔-纽曼黑洞。假定这一克尔-纽曼黑洞的外视界面积为A_1。我们知道，霍金在1971年曾证明过一个著名的定理，叫做黑洞面积定理（blackhole area theorem），它表明如果宇宙监督假设成立，且物质分布满足零能量条件，那么黑洞的视界面积永不减小〔10〕。按照这一定理，显然有$A \leq A_1$。

另一方面，克尔-纽曼黑洞的外视界面积为$A_1 = 4\pi(r_+^2 + J^2)$（其中r_+为外视界的径向坐标值，由式（4.2.1）给出）。不难证明（请读者自行验证），克尔-纽曼黑洞的质量与视界面积之间满足：$A_1 \leq 16\pi m^2$。由于$A_1 \leq 16\pi m^2$正是质量m的施瓦西黑洞的视界面积，因此这个结果的物理意义很简单，那就是在具有相同质量的所有黑洞中，施瓦西黑洞具有最大的视界面积。将这一结果与上面的不等式$A \leq A_1$联立起来，我们就得到$A \leq 16\pi m^2$。

如果我们进一步假设在整个过程中没有外来的质量（由于渐近平直超曲面S实际上就是全空间，因此这一假设显然是合理的），那么克尔-纽曼黑洞的质量显然不可能大于S上的ADM质量M，即$m \leq M$（ADM质量的定义请参阅3.3节）。将上述所有环节联系起来，我们可以看到，

如果宇宙监督假设成立，那么S上的初始条件必须满足一个不等式：$A \le 16\pi M^2$，其中A为S上所有黑洞的外视界面积之和，M为S上的ADM质量。

读者们想必认出来了，这个不等式正是我们在3.7节中介绍过的彭罗斯猜想。1973年，彭罗斯正是通过类似于我们这里所介绍的思路而提出这一猜想的。从上述思路中可以看到，彭罗斯猜想几乎可以视为宇宙监督假设的推论，或者说几乎可以视为宇宙监督假设成立的必要条件（但没有任何迹象表明它有可能是充分条件）。在4.1节中我们已经看到，宇宙监督假设远比奇点定理来得困难，而通过上面的介绍，我们又可以看到，宇宙监督假设要比彭罗斯猜想更为困难。当然，上述介绍并不构成对两者关系的数学证明，因为我们用到了一些未予严格证明的假定，比如假定所有黑洞最终都会变成克尔-纽曼黑洞。但即便如此，它仍然为寻找破坏宇宙监督假设的努力提供了一个新的可能方向，那就是寻找能破坏彭罗斯猜想的初始条件。

不过，从我们在3.7节所介绍的情况来看，虽然彭罗斯猜想目前还只是一个猜想，但人们在研究这一猜想上已经取得的进展表明，它最终被证明的可能性是很大的。相应地，通过寻找能破坏彭罗斯猜想的初始条件来构筑宇宙监督假设的反例，其希望则是很渺茫的。对于相信宇宙监督假设的人——比如霍金——来说，这显然又是一条好消息，因为彭罗斯猜想的成立虽然并不意味着宇宙监督假设一定成立，但无疑构成一种很强的支持——起码，它可以排除一大类破坏宇宙监督假设的可能性。也正因为有这么多的好消息，霍金才会信心满满地下注。只可惜，好消息虽多，却无法构成证明，而坏消息哪怕只有一条，也足以带来麻烦。

4.4　壳层穿越奇点与壳层会聚奇点

由于黑洞无毛发定理的存在，以及摧毁黑洞的努力所遭遇的困难，利用广义相对论的稳定解来寻找裸奇点的可能性起码在经典广义相对论的范围内可以说是基本被排除了[11]。不过，对裸奇点的直接寻找依然是研究宇宙监督假设的主要途径之一，只是进一步努力的方向在很大程度上转向了动态解。这其中，动态但具有良好对称性的解要比缺乏对称性的解容易研究得多，从而成为了最早获得突破的方向。

在广义相对论的所有具有良好对称性的动态解中，最简单的无疑是球对称动态解。不过我们知道，球对称引力场本身是不具有动力学自由度的——这是所谓的伯克霍夫定理（Birkhoff theorem），它表明真空中的球对称度规必定是施瓦西度规[12]。因此，非平凡的球对称动态解必须涉及物质，从而必须涉及由物质能量动量张量带来的额外复杂性。作为起步，最明智的做法显然是选择尽可能简单的能量动量张量。那么，什么样的物质具有最简单的能量动量张量呢？答案是理想尘埃物质，这种物质的能量动量张量为$T^{\mu\nu}=\rho u^{\mu}u^{\nu}$，它是理想流体能量动量张量$T^{\mu\nu}=pg^{\mu\nu}+(p+\rho)u^{\mu}u^{\nu}$在压强恒为零时的特例（由于这一缘故，理想尘埃也被称为零压理想流体），只带一个与物性有关的标量函数ρ。因此研究理想尘埃物质的球对称动态解成为了寻找裸奇点的最佳切入点之一。

另一方面，在所有动态过程中，与奇点的形成最有关系的显然是坍缩过程，因此在所有球对称动态解中，描述球对称坍缩过程的解是最值得关注的。当然，球对称坍缩过程倘若在引力的支配下进行到底，其最终产物将是施瓦西黑洞，从而是不带裸奇点的。但在坍缩的过程中，即尚处于动态的阶段中，却完全有可能出现复杂的情况。特别是，如果我们能找到一个球对称坍缩过程，使得当奇点在某个半径——包括坍缩中

心——出现时，包裹该奇点的任一同心球面之内都尚未积累起使该球面成为施瓦西视界所需的物质（从而施瓦西视界尚无法形成），则这种先于视界而出现的奇点就是裸奇点——虽然它只在有限的时间内才是裸露的〔13〕。

这样的裸奇点究竟会不会出现呢？凭空想象是得不出结果的，答案只能来自计算。这方面最早的计算比宇宙监督假设早了整整30年就已经出现了。1939年，美国物理学家奥本海默（Robert Oppenheimer，1904—1967）和他的学生施奈德（Hartland Snyder，1913—1962）研究了最简单的球对称坍缩过程：一个均匀理想尘埃球的坍缩。他们发现，这样的坍缩在随尘埃物质一同运动的观测者——即所谓的随动观测者（comoving observer）——看来，将会在有限的时间内完成，这其中奇点只有在坍缩过程彻底完成时才会出现，而这时的时空已经完全变成了施瓦西时空。不仅如此，其实早在奇点出现之前，一旦尘埃物质进入到相应的施瓦西半径以内，视界就已经形成。因此，均匀理想尘埃球的坍缩是一个视界先于奇点而形成的例子，它显然不会产生裸奇点。

看来，这又是一个有关宇宙监督假设的“利好消息”，但这一次它同时也是“利空消息”的邻居。

奥本海默与施奈德的计算如今已是每位学习广义相对论的学生都能完成的“小习题”。但广义相对论计算的困难之处就在于，哪怕是在这样的“小习题”当中加入一些额外因素，比如非均匀、有压强（即不再是尘埃物质）或有自转（即不再是球对称），问题就会急剧复杂化，甚至复杂到让奥本海默那样的人物也不得不退避三舍。这方面的研究直到30多年后才有了新的突破。1973年，德国汉堡大学的物理学家约德杰斯（Peter Yodzis）、塞弗特（Hans Seifert）等人再次对理想尘埃球的坍缩进行了研究。在他们的研究中，压强仍保持为零，物质的分布也仍维持球对称，唯一被松绑的只有均匀性条件。可以说，他们只是把潘多拉盒子打开了一条小小的缝隙。

从这条小小的缝隙里，他们会发现什么呢？

约德杰斯等人发现，当尘埃物质的密度分布满足一定的条件时[14]，不同壳层上尘埃物质的坍缩速度会出现显著差异——特别是，快速坍缩的外壳层可以穿越缓慢坍缩的内壳层。当两层尘埃物质彼此穿越时，某些物理量或几何量会出现奇异性，而此时物质的整体坍缩尚未进行到足以形成视界的程度。换句话说，在这种特殊的坍缩过程中，奇点的出现可以早于视界的形成，这意味着裸奇点出现在了坍缩过程中！约德杰斯等人的这一结果是人们最早发现的能导致裸奇点的坍缩过程之一，这种由壳层穿越而导致的奇点（实际上是球对称的奇异面）被称为壳层穿越奇点（shell-crossing singularity）。研究表明，壳层穿越奇点是一种奇异性较弱的奇点。不过，宇宙监督假设并未对奇点奇异性的强弱进行限定（我们也因此而未对奇异性的强弱作具体介绍），因此约德杰斯等人的这一结果对于宇宙监督假设具有一定的冲击性。

不过，约德杰斯等人所采用的理想尘埃模型无可避免地削弱了这一结果的现实意义。因为理想尘埃的压强恒为零，而物理直觉告诉我们，在足以形成奇点的坍缩过程中，物质将会受到高度挤压，在这种过程中压强是绝不可能为零的。有鉴于此，约德杰斯等人很快就展开并完成了一系列后续研究。在那些研究中，他们将理想尘埃换成了理想流体（从而压强不再为零），而结果则不仅维持了裸奇点的存在，还进一步证明了裸奇点在球对称微扰下具有稳定性。

除约德杰斯等人所发现的壳层穿越奇点外，人们在球对称坍缩过程中还发现了一类比壳层穿越奇点更为棘手的裸奇点：壳层会聚奇点（shell-focusing singularity）。这类裸奇点位于坍缩球体的中心[15]，只要物质分布满足特定的物态及初始条件，它既可以出现在理想流体（包括理想尘埃）的坍缩过程中，也可以出现在所谓辐射物质的坍缩过程中，后者的能量动量张量为$T^{\mu\nu}=\rho k^\mu k^\nu$（其中$k^\mu$为类光矢量）。研究表明，壳层会聚奇点的奇异性要高于壳层穿越奇点。

这样看来，在特定的物质坍缩过程中裸奇点的出现已是“铁证如山”的事实，宇宙监督假设尽管寄托了一些物理学家的良好愿望，并且也得到了不少旁证，却早在几十年前就该寿终正寝了。但事实上，这一假设却迄今仍被视为是一个尚未解决的问题，这是怎么回事呢？为什么如此直接的反例却未被认为是推翻了宇宙监督假设呢？这其中的诀窍还得到我们在4.1节中提到的那个“半拉子工程”的表述中去寻找。

4.5　走向严密表述

现在让我们回顾一下4.1节中提到的那个“半拉子工程”的表述：“在正常的物质性质及初始条件下，时空是强渐近可预测的。”我们当时就已指出，这个表述没有对“正常的物质性质及初始条件”做出明确界定，因而充其量只是一个“半拉子工程”。解决前述问题的诀窍就在这所谓“正常的物质性质”上——这个含糊其辞的条件简直就是为对付上面提到的那些“反例”而量身定做的。

在4.4节提到的两种裸奇点中，壳层穿越奇点相对来说是不足为患的，因为这种因内外壳层彼此穿越而产生的奇点在闵科夫斯基时空，甚至在非相对论流体力学中就已经存在，从而并不是广义相对论特有的。我们知道，在广义相对论之前的经典物理学中并不存在类似于奇点定理那样的结果，那里出现的奇点即便在经典物理本身的范围之内，也只不过是来自于对现实世界的过度理想化，而不具有基础意义。壳层穿越奇点既然在普通流体力学中就存在，它的起源也就昭然若揭了，那便是来自于理想流体对现实流体的过度理想化。现实流体虽然也有壳层穿越现象，但在穿越过程中是绝不会出现奇异性的。因此，出现在理想流体中的壳层穿越奇点只是理想流体这一物质模型的缺陷，而不是广义相对论的麻烦。这种本身就有问题，却想把糊涂账转嫁到广义相对论头上的物质模型，理所当然地遭到了广大爱好宇宙监督假设之士的“痛斥”，被定性为不具有“正常的物质性质”，而遭驱逐出境。

与壳层穿越奇点不同——从而更为棘手——的是，壳层会聚奇点并不总是出现在普通流体力学中，因此它起码可以部分地归因于广义相对论。但是，既然理想流体本身已经因为不具有“正常的物质性质”而遭驱逐，它所造成的一切麻烦——包括壳层会聚奇点——当然也就得一并卷铺盖了，正所谓“皮之不存，毛将焉附”，壳层会聚奇点傍错了老板，只

能怪自己命苦了。

细心的读者也许会问：辐射物质的坍缩不也可以产生壳层会聚奇点吗？难道连辐射也不具有“正常的物质性质”吗？答案是肯定的，不过不是因为辐射不具有正常的物质性质，而是因为在4.4节中被称为“辐射”的东西乃是“披着羊皮的狼”，是一种冒牌的辐射。它之所以被称为“辐射”，只不过是因为其能量动量张量$T^{\mu\nu}=\rho k^\mu k^\nu$中的$k^\mu$类光。将这种“辐射”与理想尘埃的能量动量张量相对比，可以看到它所描述的物质相当于类光的“尘埃”，这是使用“辐射”一词的核心依据。由于类光矢量也叫零矢量，因此这种“辐射”也称为零尘埃（null dust）。可惜的是，“零尘埃”虽与普通辐射一样具有类光性，其能量动量张量却并不描述任何现实世界中已知的辐射，比如电磁波（要从数学上证明这一点虽不算困难，却也不是完全轻而易举的，不过请读者想一想，有什么简单的特性可以让我们从物理上判定像零尘埃这样的“辐射”并不是像电磁波那样的真实辐射？）。因此，这种挂着羊头卖狗肉的“辐射”并不具有“正常的物质性质”。

这样一来，4.4节提到的那些裸奇点就全都被扫地出门了，“正常的物质性质”这一“无招胜有招”的含糊术语算是在关键时刻挽救了宇宙监督假设。但是，这一仗却胜得并不漂亮，因为这种一味依靠含糊术语来打游击的做法终究是不够堂正，甚至可以说是有些无赖的。我们不禁要问：说了半天，究竟什么才是“正常的物质性质”？能不能对“正常的物质性质”给出一个正面的定义，而不要总等到反例逼上门来，才事后诸葛般地从反面告诉我们什么不是“正常的物质性质”？

这是一个合情合理的问题，也是一个合情合理的要求。幸运的是，物理学家们为这份合情合理找到了一个合情合理的答案。这个答案是这样的：如果用φ_1，…，φ_n表示物质场或其分量，那么具有“正常的物质性质”的物质场被定义为满足以下形式的准线性二阶双曲运动方程组：

$$g^{\mu\nu}(x,\varphi_j,\varphi_{j;\rho})\varphi_{i;\mu\nu} = F_i(x,\varphi_j,\varphi_{j;\rho}) \qquad (4.5.1)$$

这里，准线性（quasilinear）指的是相对于最高阶（对于本方程组来说即二阶）协变导数为线性。这一定义参照了包括电磁场在内的物理学中所有基本动力学体系的主要特征，并且做了适当的推广（比如允许度规张量$g^{\mu\nu}$依赖于物质场的一阶协变导数），因而被认为具有极大的普适性[16]。更重要的是，早在1952年人们就已证明，满足式（4.5.1）的物质场在全局双曲背景时空中具有良好的初值表述（initial value formulation），即给定光滑类空柯西面上的一组初始条件（在本书中，若无特殊说明，凡初始条件都是指非奇异的），运动方程在该柯西面的一个邻域内存在唯一且连续依赖于初始条件的解。物质场所满足的这一性质保证了它在闵科夫斯基时空（那是全局双曲背景时空的最简单特例）中具有良好的动力学性质。而既然物质场在闵科夫斯基时空中具有良好性质，那么假如它与引力的耦合体系出现了问题（比如出现裸奇点），则那种问题就可以被认为是与引力的存在密切相关，而不是像壳层穿越奇点那样的“诬告”。这一特点使得上述定义很适合于宇宙监督假设的讨论。当然，对于宇宙监督假设来说，我们还要求物质场满足适当的能量条件，比如主能量条件，以排除诸如$m<0$的施瓦西时空那样的情形（$m<0$的施瓦西时空——如我们在3.4节中所说——是带有裸奇点的）。

现在我们已经接近对宇宙监督假设给出一个比较严格的表述了，不过为了完成这一表述，除“正常的物质性质”外，还必须对粗略定义中“正常的初始条件”也给出一个直接定义。这一点初看起来是不复杂的，因为考虑到黑洞和视界的定义有赖于时空的渐近平直性（参阅4.1节），所谓“正常的初始条件”，显然就应该定义为某个渐近平直类空超曲面Σ上的满足一定渐近条件的h_{ij}和K_{ij}——简单地讲，就是渐近平直的初始数据集。有关这一点，我们在讨论广义相对论的动力学时已经表述

过了，读者们可参阅3.2节，尤其是该节的注释。不过我们稍后将会看到，这“正常的初始条件”其实也是颇有玄机的。

综合上面的讨论，现在我们正式给出宇宙监督假设的一个比较严格的表述：

宇宙监督假设：在广义相对论中，如果

（1）物质场的运动方程组为形如式（4.5.1）的准线性二阶双曲方程组；

（2）物质的分布满足主能量条件；

（3）（Σ，h_{ij}，K_{ij}）为渐近平直的初始数据集。

则（Σ，h_{ij}，K_{ij}）的最大柯西展开是强渐近可预测的。

这里（Σ，h_{ij}，K_{ij}）的“最大柯西展开”（maximal Cauchy development）指的是该初始数据集所对应的解（M，$g_{\mu\nu}$）在微分同胚（diffeomorphism）意义上保持Σ为柯西面的最大延拓〔17〕。

好了，现在我们终于有了一个关于宇宙监督假设的明确表述。宇宙监督假设也已经闯过了上面介绍过的所有关卡。那么，在现在这样一个明确表述下，它还能继续过关斩将吗？

4.6　零质量标量场与裸奇点

宇宙监督假设的明确表述虽然帮助宇宙监督假设闯过了若干关卡，却很不幸地排除了诸如理想流体之类广义相对论研究者们喜闻乐见的物质分布。为了进一步研究宇宙监督假设，物理学家们不得不引进其他类型的物质。这其中最简单的莫过于标量场（scalar field），也叫做克莱因-戈登场（Klein-Gordon field）。而标量场中最简单的则是所谓的零质量标量场，它的能量动量张量为$T_{\mu\nu}=\varphi_{;\mu}\varphi_{;\nu}-(1/2)g_{\mu\nu}\varphi^{;\rho}\varphi_{;\rho}$，相应的运动方程则为最简单的线性二阶双曲方程（请读者自行证明）：

$$g^{\mu\nu}\varphi_{;\mu\nu}=0 \qquad (4.6.1)$$

这种标量场与引力耦合的体系被称为爱因斯坦-克莱因-戈登体系〔18〕，它显然满足宇宙监督假设的明确表述对运动方程的要求，因而成为了研究宇宙监督假设的新战场。

自20世纪80年代中期开始，当时任教于美国雪城大学（Syracuse University）的数学物理学家克里斯多洛（Demetrios Christodoulou，1951—　）开始对爱因斯坦-克莱因-戈登体系进行较为系统的研究。由于问题的高度复杂性，克里斯多洛不得不将注意力集中到具有球对称性的爱因斯坦-克莱因-戈登体系上。但即便引进了这么高的对称性，问题依然相当困难，以至于与宇宙监督假设有关的突破直到6年之后的1992年才由当时在美国得克萨斯大学奥斯汀分校（University of Texas at Austin）的物理学家肖普推克（Matthew Choptuik）做出。肖普推克在克里斯多洛的一些前期解析研究的基础上，对球对称爱因斯坦-克莱因-戈登体系的坍缩过程进行了数值研究。他的研究表明，如果向内坍缩的标量场的某个强度参数足够小，坍缩场将会被反射回无穷远；而如果该强度参数足够大，坍缩场则会形成黑洞。这两种极端情形都不破坏宇宙监

督假设。但是，这两种极端情形之间还存在一段参数范围，那里的情况如何呢？特别是，当强度参数由大变小，在由能够形成黑洞转变为无法形成黑洞的那一个被称为临界参数（critical parameter）的特殊数值附近的情况又如何呢？这些问题克里斯多洛曾在1987年提出过，肖普推克找到了答案，他的结果发表于1993年初。

肖普推克的数值研究表明，当强度参数由大变小接近临界参数时，黑洞的质量M_{BH}与参数p之间存在一个幂函数关系：M_{BH}～（$p-p_c$）$^{\gamma}$，这里p_c为临界参数，γ是一个对标量场具有普适性的指数，数值约为0.37。肖普推克的这个结果表明，在接近临界参数时，黑洞的质量——因而其视界——将无限缩小，外部观测者可以无限接近黑洞中心的奇点。这种可以无限接近的奇点已经非常类似于裸奇点，但还不完全是，因为对于任何给定的强度参数，只要它不完全等于临界参数，这种接近程度就终究还是有限的。那么，当强度参数完全等于临界参数时，情况又如何呢？这时候黑洞的视界将会消失，但奇点是否也会随之消失呢？如果奇点也消失了，那就仍然不存在裸奇点。幸运（或不幸——取决于看问题的立场）的是，肖普推克的研究表明，在这种临界情形下，时空的曲率标量在坍缩中心附近趋于发散，即奇点依然存在！因此，临界情形下爱因斯坦-克莱因-戈登体系的坍缩会在坍缩中心形成裸奇点。这是物理学家们首次在具有“正常的物质性质”的引力耦合体系中找到裸奇点。

比肖普推克的稍晚，印度塔塔基础研究所（Tata Institute of Fundamental Research）的物理学家德维外迪（Indresh Dwivedi）与乔希（Pankaj Joshi）给出了一个稍具一般性的结果，即在满足弱能量条件的球对称体系中，特定的非奇异初始条件可以导致裸奇点。而克里斯多洛本人也在爱因斯坦-克莱因-戈登体系的特定球对称坍缩过程中发现了性质与肖普推克所发现的临界情形相类似的裸奇点。1995年，剑桥大学的数学物理学家哈梅迪（Rufus Hamadé）与斯图尔特（John Stewart）通过

更细致的数值计算进一步证实了肖普推克的结果。

出现在零质量标量场中的这一系列结果对宇宙监督假设是沉重的打击（事实上它直接否证了我们在4.5节中所给出的宇宙监督假设的“比较严格”的表述）。虽然原则上我们仍可以有“遁词”，比如现实世界中并不存在具有基础意义的零质量标量场（粒子物理标准模型中唯一的标量场——希格斯场——是有质量的），但由于零质量标量场具备基本物质场所应具备的所有良好的动力学特征，因此我们有理由认为（虽然那不等同于证明），裸奇点在零质量标量场中的出现意味着“正常的物质性质”并不能成为挽救宇宙监督假设的救命稻草。

在这一系列的不利结果面前，久经考验的宇宙监督假设的捍卫者霍金终于扛不住了。1997年初，他——如我们在4.1节末尾所述——终于公开承认自己在与同事普雷斯基尔及索恩的赌局中落败，并依约向后者提供了赌注之一的足以覆盖对方裸体的衣服。不过，一局之败显然并不能让霍金就此俯首称臣。心有不甘的他请人在输给对方的衣服上绣上了那句著名的短语：“上帝憎恶裸奇点”，以示最终的胜负还得“走着瞧”。而且——也如我们在4.1节末尾所述——就在认输的同一天，他用一份修改了的赌约与普雷斯基尔及索恩开始了一个新的赌局。

霍金之所以要对赌约进行修改，除了因为对宇宙监督假设痴心不改外，还有一个很重要的原因，那就是他注意到人们当时针对宇宙监督假设所构造的反例全都依赖于非常特殊的初始条件，比如肖普推克的裸奇点要求标量场的强度参数恰好等于临界值，大了或是小了都不行；此外它们还大都依赖于非常严格的对称性，比如球对称性。如果宇宙监督假设的所有反例全都具有类似性质，即依赖于非常特殊的初始条件或严格的对称性，那它们在现实世界中实际上是无法出现的。有鉴于此，霍金在新赌约中规定：初始条件必须是一般的（generic）而不是特殊的。换句话说，霍金认为：一般初始条件（**generic initial condition**）下的动力学演化不会产生裸奇点〔19〕。或者反过来说：裸奇点只能由特殊的

初始条件所产生。

显然，霍金的这一附加条件可以被认为是对4.1节的粗略表述中提及过，并在上一节的“比较严格”的表述中以渐近平直初始数据集的方式有所体现——但显然体现得还不够——的“正常的初始条件”的补充。从4.5节中我们知道，“正常的物质性质”曾一度挽救过宇宙监督假设，可惜挽救得并不彻底。现在霍金把新的希望寄托在了“正常的初始条件”上，希望它能干得比“正常的物质性质”更好，甚至一劳永逸地完成“救驾”任务。由于现实世界每时每刻、每个角落都充满着各种各样的随机微扰（比如宇宙微波背景辐射、引力波等），因此霍金寄希望于从初始条件的一般性这一角度来排除裸奇点，的确是一个比较合理的想法。新赌约对赌注的规定没什么大的改变，大家虽然都不缺钱，但衣服总是需要的，因此输家仍须向赢家提供足以覆盖后者裸体的衣服。不过普雷斯基尔与索恩对霍金上次所绣的“上帝憎恶裸奇点”看来有些耿耿于怀，因此在新赌约中特意强调衣服上所绣的必须是一句“真正愿赌服输”（truly concessionary）的话，而不能再玩花招。

4.7 讨论

霍金等人的新赌约问世后不久，1999年，上文提到的克里斯多洛在普林斯顿大学及高等研究院（Institute for Advanced Study）主办的《数学年报》（Annals of Mathematics）上发表了一篇文章（这篇文章的基本结果其实早在霍金等人的新赌约问世之前就已完成，并曾被圈内人士广泛引用）。在这篇文章中，他对爱因斯坦-克莱因-戈登体系做了一些解析研究，目的是想判定形成裸奇点的初始条件是否稳定。他的研究表明，对于爱因斯坦-克莱因-戈登体系来说，即便在球对称的情况下，能形成裸奇点的初始条件也是不稳定的，它会被特定类型的球对称微扰所破坏。显然，克里斯多洛的这一结果对霍金的新赌约是一个很大的支持，尽管它距离证明“霍金版”的宇宙监督假设无疑还差得很远。

霍金等人的新赌局迄今尚未见出输赢。这份赌约虽不是学术作品，却在很大程度上指出了宇宙监督假设的一个重要的研究方向。因为事实上，种种反例的存在已经推翻了如我们在4.6节中所表述的那种宇宙监督假设。要想挽救宇宙监督假设，附加条件的引进已是势在必行，而霍金在新赌约中引进的初始条件的一般性，在所有有可能挽救宇宙监督假设的附加条件中，几乎是唯一一种不至于将之弱化到失去现实意义的条件。不过，尽管霍金引进的附加条件不至于将宇宙监督假设弱化到失去现实意义，但却仍有可能造成一些不容忽视的后果。比如我们知道，广义相对论中有一些重要的结果——比如我们在4.3节中提到过的黑洞面积定理——是在类似于4.6节所给出的那种如今看来并不成立的宇宙监督假设的基础上证明的。霍金引进的附加条件对那些证明及其结论会产生什么影响？这也许是值得有志于广义相对论研究的读者去思考的。

在结束本专题之前，让我们回过头来提一下4.1节留下的一个伏笔。读者们也许还记得，我们在给出宇宙监督假设的粗略版时曾经提

到，那是所谓的弱宇宙监督假设。什么是弱宇宙监督假设？它是指只要求奇点存在于视界以内，从而与视界外的时空断绝因果联系，却并不排除视界内的观测者直接观测到奇点的可能性。但是，假如我们认为经典广义相对论必须自洽到不允许裸奇点具有观测意义的程度（这是人们提出宇宙监督假设的一个重要动机），那么视界内的观测者似乎也不应该例外。有人可能会说：视界内的观测者既然再也不可能回到视界之外了，那无论他们观测到什么，都已不再具有实际意义。这话初听起来不无道理，但我们别忘了，黑洞原则上是可以任意大的，可以大到不仅让观测者能平安穿越视界（黑洞越大，观测者穿越视界时所经受的潮汐力就越小），而且还可以让他们在黑洞内存在任意长的时间，甚至可以在那里从事系统的物理研究。在那里难道他们会发现完全不同的物理学吗？对于这样假想性的问题，不同的人显然会有不同的看法。但有一些物理学家相信，既然我们是由于对经典广义相对论的自洽性具有信心而提出了宇宙监督假设，就没有理由把视界内的观测者排除在外。他们提出，裸奇点不仅对于视界外的观测者，而且对于视界内的观测者都是不具有观测意义的。这就是所谓的强宇宙监督假设。

依据强宇宙监督假设，只要排除像大爆炸那样的初始奇点，时空的所有区域都应当能被适当的初始条件所预测，因此强宇宙监督假设曾被彭罗斯表述为：除去初始奇点外，所有物理上合理的时空都是全局双曲的。当然，对强宇宙监督假设的严格表述将会涉及一些数学上的微妙细节，以排除通过简单拓扑剪拼手段构造出的反例，这些我们就不在这里介绍了。

与弱宇宙监督假设的初看起来正确不同，强宇宙监督假设似乎初看起来就不太正确。我们早已知道，奇点在经典广义相对论中是无可避免的。由于强宇宙监督要求奇点对所有观测者都不具有观测意义，这意味着仅仅像弱宇宙监督假设所要求的那样用视界包裹奇点是不够的，为了使视界内的观测者也无法观测到奇点，奇点必须是类空的。类空奇点

的一个例子是施瓦西奇点，观测者虽然可以依据广义相对论预知其存在，但只有在真正撞上时才能“观测”到这样的奇点。但是，并非所有奇点都是类空的，比如由雷斯勒-诺斯特朗姆或更一般的克尔-纽曼度规描述的奇点就是类时的。对于那样的奇点，所有进入内视界的观测者都有可能直接观测到它的存在及可能的物理效应。

读者也许会觉得奇怪，既然像雷斯勒-诺斯特朗姆和克尔-纽曼这样重要的时空就已经允许进入内视界的观测者直接观测到奇点，从而违反强宇宙监督假设，怎么还会有人“不识时务”地提出强宇宙监督假设呢？这是因为有迹象表明像雷斯勒-诺斯特朗姆和克尔-纽曼时空中的内视界这样的东西很可能是不稳定的。这些内视界是所谓的柯西视界，它是强宇宙监督假设遭到破坏时必然会出现的[20]。1982年，钱德拉塞卡（Subrahmanyan Chandrasekhar，1910—1995）与哈特尔（James Hartle，1939— ）证明了雷斯勒-诺斯特朗姆时空的柯西视界是一个无限蓝移面，即来自远处的辐射在这一视界处会被无限蓝移，从而不仅会对试图穿越这一视界的任何观测者造成致命伤害，而且会严重干扰雷斯勒-诺斯特朗姆度规本身，导致柯西视界的不稳定。很多物理学家认为，类似的不稳定性也存在于更一般的克尔-纽曼时空中，其结果便是使得像诸如雷斯勒-诺斯特朗姆或克尔-纽曼度规所描述的那种奇点为类时的情形无法出现[21]。当然，这些基于特例所做的研究距离证明强宇宙监督假设所需的普遍性还相差很远。

有关宇宙监督假设的研究还在进行。就笔者个人的看法而言，虽然我对霍金版的宇宙监督假设是否成立这样的纯技术问题并无先入之见，不过我并不看好导致宇宙监督假设的那种推理。物理学家们大都相信，奇点在广义相对论中的存在是广义相对论作为经典理论的局限之处。在一个完整的量子引力或更宏伟的所谓终极理论（theory of everything）中，那样的奇点是不会存在的。既然如此，我们有什么理由认为广义相对论需要以某种精巧的方式阻止事实上根本就不存在的奇点的观测效应

呢？就拿霍金那句“上帝憎恶裸奇点”来说吧，如果“上帝”创造的原本就是一个量子引力世界，在那里原本就不存在真正的奇点，那它又有什么理由要在广义相对论这样一个非量子、从而注定是片面的理论中“憎恶”一个在完整理论中根本就不存在的裸奇点呢？退一步讲，倘若“上帝”果真“憎恶”广义相对论中的裸奇点，为何又允许它出现在种种个例之中呢？导致宇宙监督假设的那种推理，无论是作为拟人化的“上帝”的喜好，还是作为单纯的物理分析，都显得不够自洽。从历史上看，在黑洞概念问世之初，英国物理学家爱丁顿曾认为大自然将会阻止黑洞出现，就像宇宙监督假设的支持者们认为大自然将会阻止裸奇点出现一样，结果他错了；在量子力学问世之初，爱因斯坦曾像霍金那样猜测过“上帝”的喜好，认为他老人家不会掷骰子，结果他也错了。历史将会在宇宙监督假设这里重演一遍，还是玩出新的花样？宇宙监督假设的最终命运将会如何？这是未来广义相对论研究中一个引人注目的悬念。

注释

〔1〕如果宇宙在未来演化中出现整体性的大坍缩（从目前所知的宇宙学常数来看，这应该不会发生），那么大坍缩也将是一个裸奇点。不过这个裸奇点是未来奇点，也是与可观测宇宙有因果关联的时间本身的终结，算不上破坏广义相对论的预言能力。

〔2〕不过，即便在非渐近平直时空中，只要存在观测意义上足够接近渐近平直条件的时空区域，黑洞的概念就仍可近似成立。但初学者常为之困惑的“宇宙本身是否是黑洞”的答案则是否定的。因为对于宇宙本身来说，哪怕在近似意义上也不存在定义黑洞所需的渐近平直区域，因此黑洞这一概念并不适用。

〔3〕有些文献——比如沃尔德的著名教材《广义相对论》（General Relativity）——在某些表述中将渐近平直与强渐近可预测列为彼此独立的条件。但实际上强渐近可预测时空的定义（包括沃尔德采用的定义）往往首先就假定了渐近平直性。

〔4〕强渐近可预测时空定义的严格表述会涉及一些微妙的数学细节，其中包括我们在3.1节中定义渐近平直时空时引进的“非物理时空”。此外，不同文献采用的定义有时带有细微差

别，这些我们就不展开讨论了。

〔5〕从理论上讲，奇点既可以存在于黑洞内部，也可能存在于视界上（可以证明黑洞是闭区域，视界是黑洞的一部分）。后一种可能性通常被忽略（即假定视界是非奇异的），不过对于宇宙监督假设来说，我们只需强调时空的强渐近可预测性，奇点是否出现在视界上对讨论没有影响。

〔6〕克尔-纽曼度规有时也被简称为克尔度规。不过克尔度规有时又专指1963年克尔发现的不带电的稳定轴对称真空解。对于这些命名上的细微差异，读者在阅读文献时需按照上下文进行判断。

〔7〕需要注意的是，在这些考虑中引力辐射及入射粒子对背景时空的干扰通常被忽略了。此外，文献中似乎未见有人分析过向极端带电黑洞投放带角动量的粒子，或向带有极端角动量的黑洞投放电荷的情形。沃尔德曾发表过一篇题为“摧毁黑洞的理想实验”（Gedanken Experiments to Destroya Black Hole）的文章，对向极端克尔-纽曼黑洞投放电荷的情形进行系统分析。但他的分析看似普遍，却并不适用于黑洞带极端角动量（即J＝m，Q＝0）的特殊情形。另外值得一提的是，对背景时空为渐近de Sitter时空的情形所做的类似研究表明，在那种情形下通过向黑洞投放电荷或角动量来摧毁黑洞是可以做到的。不过这对我们的讨论并不产生影响，因为黑洞的定义、时空的全局性质等，本身就与背景时空的渐近结构有着密切关系，不同的背景时空所对应的物理问题是完全不同的。

〔8〕这一结果与所谓的黑洞热力学第三定律也有很好的相容性，因为按照黑洞热力学，极端黑洞的温度为零，而黑洞热力学第三定律表明无法使黑洞的温度达到绝对零度，即无法使黑洞达到极端状态。

〔9〕彭罗斯最初考虑的是球对称的辐射物质，也叫做零尘埃（null dust）——即能量动量张量为$\rho k^{\mu}k^{\nu}$（k^{μ}为类光矢量）的物质。但球对称与辐射物质这两个条件对论述来说并不是必需的。

〔10〕有些文献将证明黑洞面积定理所需的能量条件表述为弱能量条件，甚至强能量条件，这是不必要的要求。

〔11〕只能说是“基本被排除”而不是彻底被排除，因为并不存在关于这一排除的严格证明，而且直到近期仍有个别物理学家在进行这方面的努力。

〔12〕伯克霍夫定理的一种早期，并且直到现在依然很流行的表述形式是“真空中的球对称度规必定是静态度规”。不过后来人们发现施瓦西度规在延拓区域内不一定是静态的，因此“真空中的球对称度规必定是施瓦西度规”被认为是更准确的表述。

〔13〕严格地讲，使该奇点成为裸奇点还要求当施瓦西视界形成时，从奇点发出的光信号已经传播到视界之外（请读者想一想，这是为什么？）。

〔14〕那些条件全都是在满足适当的能量条件——比如主能量条件——的基础之上提出的。

〔15〕壳层会聚奇点既可以位于有限半径处，也可以位于坍缩球体的中心，不过位于有限半径处的壳层会聚奇点通常晚于视界而出现，从而不是裸奇点。

〔16〕请读者想一想，理想流体为什么不满足这一定义？

〔17〕柯西面的定义可参看2.4节，“最大柯西展开”的详细定义可参阅沃尔德的教材。

〔18〕爱因斯坦-克莱因-戈登体系中的标量场一般来说是有质量的，本节涉及的零质量情形只是特例。

〔19〕霍金新赌约中涉及“一般初始条件”的原文是“源自一般初始条件（即源自初始条件组成的开集）的动力学演化不可能产生裸奇点”。这一表述似乎只有从“能产生裸奇点的初始条件集不包含开子集”这一角度来理解才既有意义，又不与已知的反例相矛盾。如果能产生裸奇点的初始条件集不包含开子集，那么任何能产生裸奇点的初始条件都可以被某些微扰所破坏（请读者结合开集的定义自行论述这一点），从而是不稳定的。但另一方面，经过这样理解的“一般初始条件”并不意味着能产生裸奇点的初始条件是“稀有”的（请读者想一想，这是为什么？）。从而严格讲并不足以剥夺裸奇点的现实意义。可惜霍金似乎从未在学术论文中阐述过这一条件，从而使我们无法找到更可靠的参照。在笔者看来，对“一般初始条件”的更恰当的定义是要求能导致裸奇点的初始条件的相对测度为零。在这方面，宇宙监督假设的另一位核心支持者彭罗斯的看法要狡猾得多，他认为现阶段最好别给“一般初始条件”下明确定义，以便保留随机应变的能力。

〔20〕这是因为强宇宙监督假设要求时空具有全局双曲性，因而它的破坏意味着初始条件的柯西展开具有边界，这一边界——如我们在2.4节中所介绍的——正是柯西视界。

〔21〕需要注意的是，柯西视界在赖斯纳-诺斯特朗姆和克尔-纽曼情形下的不稳定性并不

意味着它在任何情形下都不稳定。一个引人注目的反例是：莫里斯（Mike Morris）、索恩和尤瑟福（Ulvi Yurtsever）曾在1988年发现，虫洞时空中的柯西视界不具有由无限蓝移导致的不稳定性。不过，如我们在虫洞物理学中将会看到的，虫洞时空并不满足宇宙监督假设所要求的能量条件。

第5章 虫洞物理学

5.1　萨根的小说

从本节起，我们将进入本书的最后一个专题：虫洞物理学。虫洞（wormhole）这个朴素的术语会让人产生一些朴素的联想，比如联想到虫子在苹果上咬出的洞。它们在拓扑学上的含义也确实类似，只不过那“苹果”要换成时空，而那“虫子”有可能是我们自己。如果联想得更多一点的话，还可能由拓扑而联想到另一个层面上，那就是度量，或者说长度。苹果中的虫洞若足够直的话，是可以成为连接苹果表面两个点的捷径的——即比沿苹果表面的路径有更短的长度。物理学家希望时空中的虫洞也能如此，即成为连接两个时空区域的捷径。

那样的捷径对于时常要用自己的想象力去触碰光速极限的科幻作家来说，无疑是便利的工具和巨大的福音。不过，与科幻作家从科学概念中寻找工具的老套故事不同的是，在虫洞这一工具的问世过程中，科幻作家本身就起了一定的促进作用，这一点是足可令科幻作家们自豪的。唯一使这自豪稍稍逊色的，是那起了促进作用的科幻作家乃是业余的，或者说只是一位客串者——他的正式身份是天文学家。

1985年，美国康奈尔大学（Cornell University）的天文学家萨根（Carl Sagan，1934—1996）发表了一部很著名、并在12年后（即1997年）被拍成电影的科幻小说，书名叫做《接触》（Contact）。在这部小说中，包括女主人公在内的五位地球人（在电影中改为了女主人公一人）通过以外星智慧生物传来的信息为蓝图建造起来的装置，在无论对于他们自己还是对于其他人来说都很短的时间内，旅行到了距地球26光年的织女星（Vega）附近，与外星文明进行了接触，并顺利返回了地

球。

小说中的人物是如何跨越广袤的恒星际空间的呢？聪明的萨根让他们自己进行了讨论，讨论的结果是认为他们的旅行是通过虫洞进行的。萨根并且采用了我们上面提到的有关苹果的比喻，表示——当然，仅限于他的小说之中——宇宙就像一个大苹果，我们人类生活在它的表面上，而外星智慧生物则建造了穿越苹果内部的纵横交错的虫洞，作为星际旅行的通道。

虽然只是科幻世界的客串者，但萨根所具有的天文学家与科普作家的双重身份，以及此前不久因电视科普系列片《宇宙》（Cosmos）而获得的公众知名度，使他的小说产生了巨大影响。事实上，他的小说尚未动笔就已获得了200万美元的据说是史上最高的预付稿费。而小说发表之后，他所采用的虫洞这一概念，则很快风靡科幻界，成为了其他科幻作家竞相追随的标准概念，以及星际旅行的标准手段。

在撰写小说的那段时间里，萨根住在纽约上州一栋颇具乡村风格的屋子里。周围是山丘地带，地势起伏不定，还间杂了峡谷、沟堑、悬崖、瀑布等。美国科普作家哈尔彭（Paul Halpern）在一本有关虫洞的科普作品中曾饶有兴致地罗列了那些地形，并风趣地表示那些都是诱使人们思考时空结构的理想地形。确实，在那样复杂多变的地形下，人们很少能沿直线从一处旅行到另一处，要想有快捷的旅行，就得在某些地方构建像虫洞那样的通道。不过，地形虽然理想，萨根恐怕却并未从中得到过有关时空结构的灵感，起码他小说中的虫洞概念并非受地形启发，而是另有缘由的。事实上，尽管萨根对于虫洞概念的流行甚至兴起都有着促进之功，最主要的功劳却必须归于他的一位老朋友——加州理工学院（California Institute of Technology，Caltech）的物理学家索恩。

那是虫洞物理学中的一段佳话，索恩本人曾就此写过一段回忆：

我刚刚教完1984—1985学年的最后一堂课，沉陷在办公室的椅子上让激情退去，这时电话

铃声响了起来。打来电话的是我的老朋友、康奈尔大学的天体物理学家萨根。“很抱歉打扰你，索恩。”他说：“我刚刚写完一部有关人类与外星文明首次接触的小说，但却有些顾虑。我希望其中的科学尽可能地准确，可我担心自己也许会搞错某些引力理论的东西。你愿意看一下并给我一些建议吗？”我当然愿意。那将是有趣的，甚至有可能是好玩的，因为萨根是个聪明的家伙。更何况，我怎能拒绝这种来自朋友的要求呢？

几星期后，萨根寄来了一叠厚厚的书稿。当时索恩正要与前妻及儿子驱车前往女儿就读的大学参加毕业典礼，便带上了书稿。一路上，索恩的前妻和儿子轮流开车，索恩自己则将书稿浏览了一遍。“小说很有趣，但萨根确实遇到了麻烦。”——这是索恩读完书稿后的感觉。为什么会有这种感觉呢？因为萨根在书稿中所设想的星际旅行是通过黑洞进行的，而在索恩这样的引力理论专家眼里那是一条死路。在参加完毕业典礼返家的途中，当车子行驶在平坦的州际公路上时，索恩开始为萨根出谋划策。经过一番思考，他想到了十几年前由他的导师、美国物理学家惠勒提出的虫洞概念〔1〕。这一概念当时还知者甚少，但索恩无疑是例外。这一概念可以帮助萨根吗？通过虫洞有可能进行星际旅行吗？索恩当即拿出理论物理学家的看家“法宝”——纸和笔——进行了分析，并得到了一个喜忧参半的初步结果（我们将在后文中详细介绍）。尽管只是喜忧参半，比起黑洞那死路一条来终究是乐观多了。回到家后，索恩便将通过虫洞进行星际旅行的设想及初步的理论分析告诉了萨根，后者接受了他的设想，于是就有了出现在《接触》这部小说、后来的同名影片，以及许许多多其他科幻作品中的那些通过虫洞进行星际旅行的精彩故事。

在向萨根推荐了通过虫洞进行星际旅行的设想之后，索恩自己也对虫洞产生了浓厚兴趣，与他的学生莫里斯一起，开始对虫洞物理学展开认真研究，并于两年后（即1987年）发布了结果。索恩和莫里斯的研究是对虫洞结构的第一次细致研究（那之前的虫洞基本上只是一个概

念），从而堪称是虫洞物理学的开山之作。不过，索恩和莫里斯自己对他们的研究做了一个很谦虚的定位，只定位为讲授广义相对论的教学工具。这可以从他们论文的标题中看出来，那标题是“时空中的虫洞及它们在星际旅行中的应用：讲授广义相对论的工具”（Wormhole in spacetime and their use for interstellar travel: A tool for teaching general relativity）。与这一定位相一致的是他们为论文所选的刊物：《美国物理杂志》（American Journal of Physics），那是一份著名的物理教学刊物。

5.2 黑洞、白洞和虫洞

不过，在介绍索恩和莫里斯的研究之前，让我们先对使索恩觉得萨根“确实遇到了麻烦”的设想——即通过黑洞进行星际旅行的设想——作一个简单的介绍及评述。这不仅是因为对一个“此路不通”的设想的正式排除，更能显出虫洞的意义，以及它在现代科幻作品中快速风靡的原因，而且也是因为那“此路不通”的设想遭遇失败的地方，恰恰孕育着虫洞物理学崛起的关键，从而对那些失败的了解，有助于我们更好地理解虫洞物理学。

初看起来，通过像黑洞那样恶名远扬，被认为能吞噬一切（其中当然包括宇航员）的危险天体来进行星际旅行是一种离奇的设想。不过，结论不能下得太快，因为萨根毕竟不是普通的科幻作家，能被他采用的设想除了离奇之外，多少是有些理论渊源的。事实上，那设想不仅有理论渊源，而且那渊源还非同一般，因为它来自于广义相对论的创始人——爱因斯坦。

1935年，爱因斯坦与以色列物理学家罗森（Nathan Rosen，1909—1995）合作发表了一篇题为“广义相对论中的粒子问题”（The Particle Problem in the General Theory of Relativity）的论文。在那篇论文中，爱因斯坦和罗森提出了后来被称为爱因斯坦-罗森桥（Einstein-Rosen bridge）的广义相对论的特殊解。

什么是爱因斯坦-罗森桥呢？它有若干个版本，其中最简单的版本是对施瓦西度规

$$ds^2 = \left(1 - \frac{2m}{r}\right)dt^2 - \left(1 - \frac{2m}{r}\right)^{-1}dr^2 - r^2 d\Omega^2 \qquad (5.2.1)$$

作坐标变换$r-2m=u^2$所得到的解（请读者自行验证）：

$$ds^2 = \frac{u^2}{u^2 + 2m}dt^2 - 4(u^2 + 2m)du^2 - (u^2 + 2m)^2 d\Omega^2 \quad (5.2.2)$$

这个解在u∈（-∞，∞）的整个取值范围内都不包含r＜2m的区域（从而避免了r＝0的奇点），且在u→-∞和u→∞时都是渐近平直的（因为对应于r→∞），从而像是一座连接u→-∞和u→∞这两个渐近平直时空的“桥梁”。既然有“桥梁”，自然可以联想到旅行。不过，爱因斯坦和罗森提出这个解的用意与星际旅行是毫不相干的，他们的用意——如他们文章的标题所明示的——乃是用场的概念来表示粒子。这种试图将场和粒子统一起来（确切地说是统一在场的框架内）的努力是爱因斯坦晚年研究的核心课题之一，也是他寻求统一理论的一个重要环节。上述爱因斯坦-罗森桥在爱因斯坦眼里乃是广义相对论的一个处处有限（因为避免了r＝0的奇点）、有可能用来表示中性粒子的特殊解。

爱因斯坦的这一努力与他晚年在寻求统一理论之路上的其他尝试一样，并未取得成功。但爱因斯坦-罗森桥所具有的连接两个渐近平直时空的特性，起码在表观上与虫洞有着异曲同工之处[2]。由于上述爱因斯坦-罗森桥是从描述施瓦西黑洞的施瓦西度规中构造出来的，因此可在一定程度上被视为是通过黑洞进行星际旅行这一设想的理论渊源。

不过，表观上的异曲同工并不等于实质上相似，渊源的非同一般也并不意味着设想可信。如果你试图——像萨根原先设想的那样——通过上述爱因斯坦-罗森桥进行星际旅行，你的结局将会是很悲惨的，因为你将不会经过任何“桥梁”，更不会到达什么渐近平直时空，而是会直接落入施瓦西黑洞。爱因斯坦和罗森通过本质上是将黑洞外部区域覆盖两次的特殊坐标变换构造出的所谓“桥梁”，其实只是坐标缺陷带来的幻象，因为在施瓦西视界（r＝2m或u＝0）处，爱因斯坦-罗森桥的度规是退化的（因为g_{00}＝0）。除上述最简单的版本外，爱因斯坦-罗森桥还有更复杂的版本，被爱因斯坦和罗森用来试图表示带电粒子。那种版本的问题更多，除同样存在坐标缺陷外，还有其他方面的问题，就不赘述

了。

而更糟糕的是，即便“视死如归”地撇开上述所有问题，通过爱因斯坦-罗森桥进行星际旅行也还存在其他麻烦，比如众所周知（感兴趣的读者可以自行验证），物体抵达施瓦西视界（即抵达上述爱因斯坦-罗森桥的正中央）的过程在外部观测者看来是要花费无穷长时间的。这使得通过上述爱因斯坦-罗森桥进行星际旅行即便可以“苟全性命”，也只对旅行者本人才有意义，外部观测者却是永远也无法看见旅行过程的完成，更不可能从中受益（哪怕是信息意义上的受益）的。

因此，爱因斯坦-罗森桥是不能够作为星际旅行的通道的。

不过，这并不等于通过黑洞进行星际旅行就被排除了，因为爱因斯坦-罗森桥只是通过黑洞进行星际旅行的若干设想之一。这方面的另一类设想，是通过广义相对论的某些更复杂的解（比如雷斯勒-诺斯特朗姆解或克尔-纽曼解）的某些奇妙特点（比如奇点成为类时甚至变成奇环，从而可被绕开或穿越）进行星际旅行。但是，这类设想也面临一些棘手的问题，比如我们在4.7节中曾经提到过，雷斯勒-诺斯特朗姆解和克尔-纽曼解的内视界很可能是所谓的无限蓝移面，它会对任何试图穿越或靠近的旅行者造成致命伤害，甚至破坏雷斯勒-诺斯特朗姆解和克尔-纽曼解本身。此外，黑洞引力场的不均匀性所带来的所谓潮汐力（tidal force）也很可能会对旅行者产生致命伤害〔3〕。

除上述种种麻烦外，通过黑洞进行星际旅行的设想还面临另外一个棘手的问题，那就是出口问题。黑洞作为吞噬一切的家伙，进入它不是问题，但进入之后不管从什么地方出来，都似乎有悖其“吞噬一切”的恶名。这个问题该如何解决呢？这就引出了一个或许算得上是引力理论中最离奇的概念——白洞（white hole）。

顾名思义，白洞就是性质与黑洞完全相反的天体：黑洞只进不出，白洞就只出不进；黑洞吞噬一切，白洞就喷射一切。那样的天体如果存在，就有可能作为通过黑洞进行星际旅行时的出口。但与黑洞之顺应引

力的基本性质，从而受到理论家们的普遍青睐，甚至获得了较强的间接观测支持不同，白洞的喷射乃是逆引力而行，从而依赖于极为离奇的初始条件。事实上，白洞这一概念之所以存在，唯一的理由就是作为黑洞的时间反演——因为广义相对论具有时间反演对称性。但我们都知道，很多原则上可以存在的时间反演过程在实际上是几乎绝不可能出现的。比如你扔一块石头到水里，发出“扑通”的声音，并荡起一圈圈向外延展的涟漪，这个过程的时间反演，即涟漪和声波向内收缩，将石头从水里反弹到你手里的过程，虽然原则上是可以存在的，但实际上却几乎绝不可能出现。白洞作为黑洞的时间反演也是如此。广义相对论的黑洞解所具有的极大的简单性也许会给人一个错觉，以为黑洞的时间反演未必罕见，但实际上，若果真要反演一个黑洞，还必须反演它的蒸发过程，仅此一点，就足以将概率降低到难以想象的小。更何况，倘若要反演物体落入黑洞的过程，就还必须反演这一过程中发射的各种辐射（包括引力辐射）等，概率之小就更是无法形容了。不仅如此，所有跟黑洞有关的麻烦，在白洞上也会有一定程度的体现。比如拿最简单的施瓦西黑洞来说，物体抵达视界的过程在外部观测者看来需要无穷长的时间，相应的，假如有白洞的话，物体从视界中喷射出来的过程在外部观测者看来也将需要无穷长的时间，或者换句话说，喷射过程在外部观测者看来应该是从无穷长时间之前就开始了，这在可观测宇宙中是根本不可能实现的。因此，白洞的背后即便不是彻底的死路，起码也是希望极为渺茫的，顺带着也给通过黑洞进行星际旅行的设想又泼了一大瓢冷水。

黑洞、白洞都靠边站了，自然就轮到虫洞出场了[4]。虫洞这一概念——如前所述——乃是出自索恩的导师、美国物理学家惠勒之手。20世纪50年代，惠勒展开了对所谓“几何动力学”（geometrodynamics）这一新兴、可惜始终未能真正兴起的领域的研究。这一研究的核心目标是将物理学几何化。为了达到这一目标，一个必不可少的步骤就是将粒子（包括带电粒子）几何化。细心的读者也许看出来了，惠勒的这一目标

跟爱因斯坦和罗森研究爱因斯坦-罗森桥的用意十分相似。事实上，上面所介绍的最简单的爱因斯坦-罗森桥正是试图用已经几何化了的引力场来表示粒子，从而可以视为将粒子几何化的早期尝试。而惠勒的工作则是继爱因斯坦-罗森桥之后，这一领域内20年间几乎唯一重要的新尝试。

在这种新尝试中，惠勒引进了物理空间的多连通结构，这种多连通结构的一个简单图示出现在他发表于1955年的题为“几何子”（Geons）的论文中，正是典型的虫洞图示（见图5.1）。在这一图示中，惠勒画出了一些电磁场的力线，它们从虫洞的一个“嘴巴”（mouth）进入，又从另一个“嘴巴”出去，无须任何电荷就显示出了如同一对正负电荷所产生的电磁场。两年后（即1957年），惠勒与自己的学生米斯纳在一篇题为“作为几何学的经典物理学”（Classical physics as geometry）的论文中，将上述无须任何电荷就能显示出如同一对正负电荷所产生的电磁场的巧妙结果称为“没有电荷的电荷”（charge without charge）。类似地，他还提出了“没有质量的质量”（mass without mass），以及“没有场方程的场方程”（field equations without field equations）等颇带哲理气息的概念。这些都是他那几何动力学中的主要结果。除这些结果外，惠勒与米斯纳的那篇1957年的论文，也是虫洞这一术语的诞生地。惠勒与米斯纳在文章中表示，拓扑学家们会把他们提出的结构称为“多连通空间”（multiply-connected space），但物理学家们也许可以将之更生动地命名为虫洞——这正是索恩所继承，并经萨根普及出去的命名。

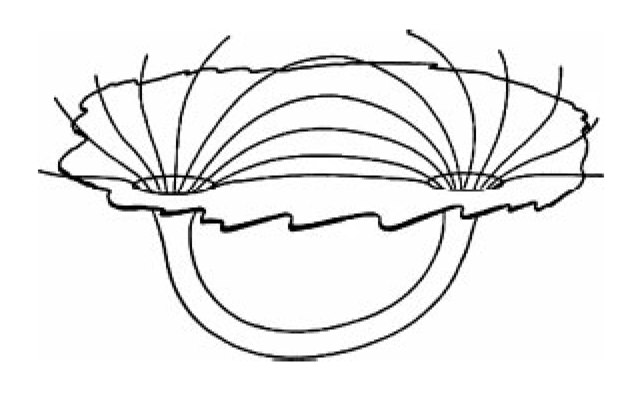


图**5.1**　惠勒的虫洞

与上述想法相匹配，惠勒还提出了有关时空结构的量子涨落（quantum fluctuation）的想法。按照这种想法，时空本身也像其他物理体系一样，在小尺度上会出现量子涨落。在极端情形下，比如当尺度小于所谓的“普朗克长度”（Planck length）时[5]，量子涨落将会变得如此巨大，就连时空的拓扑结构都有可能被改变，甚至有可能裂成碎片，出现所谓的“时空泡沫”（spacetime foam）。如果作个比喻的话，那么时空就像海面，从大尺度上看平滑如镜，随着尺度的缩小渐渐显出起伏，当尺度小到一定程度时，则可以看到汹涌的波涛乃至飞散的泡沫。惠勒的这种想法与包括虫洞在内的他那几何动力学中的若干基本概念是密切相关的。事实上，时空结构的这种量子涨落——如果存在的话——无疑可以成为虫洞出现的温床。

不过，惠勒所设想的时空结构的量子涨落从单纯类比的角度看虽颇为自然，并且受到了物理学家们迄今尚未建立量子引力理论这一不幸事实的自动“保护”，但许多物理学家对之仍是深深怀疑的，对此将在后文中作进一步介绍。另一方面，惠勒的上述想法所涉及的包括虫洞在内的各种时空结构基本上都是微观的，这对于他所追求的将粒子几何化的目标来说或许是足够了，但对于我们——以及萨根和索恩等人——感兴趣的星际旅行的需要来说却无疑差得很远。而且惠勒虽然提出了虫洞概念，但基本上只停留在概念层面，而并未对诸如虫洞的稳定性之类的技术细节给予足够关注，可那些技术细节对于通过虫洞进行星际旅行来说，却有着举足轻重的重要性。那些技术细节正是索恩和莫里斯的研究重点。因此，惠勒虽然是虫洞研究的先驱者，对虫洞物理学进行细致研究的荣誉却要归于索恩和莫里斯。索恩和莫里斯所研究的可作为星际旅行通道的虫洞有一个专门的名称，叫做“可穿越虫洞”（traversable wormhole）。索恩和莫里斯是可穿越虫洞研究的开创者。

5.3 球对称可穿越虫洞

可穿越虫洞与惠勒提出的概念层面上的虫洞的差别就在于“可穿越”三个字。究竟什么样的虫洞是可穿越的呢？这是索恩和莫里斯首先要确定的。我们在上节中介绍通过黑洞进行星际旅行的设想时曾经表示，该设想遭遇失败的地方，恰恰孕育着虫洞物理学崛起的关键。现在就让我们来盘点一下通过黑洞进行星际旅行的失败之处。

通过黑洞进行星际旅行的最核心的失败之处显然就在于存在视界，它所导致的困难是多重的：比如它的“只进不出”要靠白洞那样的离奇概念来“解套”；比如落入或离开它的过程在外部观测者看来要花费无穷长的时间；比如它有可能是致命的无限蓝移面。因此，可穿越虫洞所需满足的首要条件就是不存在视界。

通过黑洞进行星际旅行的另一个失败之处是引力场的不均匀性造成的潮汐力，虽然如我们在上节的注释中所说，这个失败之处并没有通常渲染的那样严峻，但它无疑是可穿越虫洞必须“引以为戒”的。因此，可穿越虫洞所需满足的另一个条件是穿越过程中遇到的潮汐力是人体能够承受的。考虑到潮汐力未必是穿越过程中有可能遇到的唯一应力，更普遍的条件可以表述为穿越过程中遇到的应力是人体能够承受的。

这两条就是从通过黑洞进行星际旅行的失败中得到的“经验教训”。除此之外，可穿越虫洞还必须满足一些一般性的理论条件：首先是它必须满足广义相对论场方程；其次是它的物质分布必须是物理上可以实现的——这包括物质的能量动量张量是物理上存在的，以及物质的数量是可观测宇宙可以提供的；最后则是它必须能在微扰下保持稳定——否则的话，星际飞船通过时带来的干扰就有可能破坏可穿越虫洞。

这些就是索恩和莫里斯归纳出的可穿越虫洞所需满足的条件。不过这些条件都很一般，为了便于具体计算，他们还引进了一些简化条件：

首先是假设了可穿越虫洞的度规是静态球对称的。这当然不是必需的，但在广义相对论研究中乃是首选的简化条件，比如广义相对论的第一个严格解——施瓦西解——就是在这一简化条件下得到的，从它入手进行可穿越虫洞研究也是顺理成章的。而且从物理上讲，虫洞如果是一种大尺度物质结构，它的天然形态也确实有可能像其他大型天体一样是接近静态球对称的。其次是假设了可穿越虫洞的所谓“喉咙”（throat）——即径向坐标值的最小处——是唯一的，或者换句话说，径向坐标r作为径向本征距离$s=\int dr$的函数有唯一的最小值r_0。这当然也不是必需的，因为虫洞的“喉咙”完全可以是更复杂的。不过如我们将会看到的，“喉咙”是虫洞性质最独特的地方，因此对它的简化是很有帮助的。最后则是假设了可穿越虫洞的出入口分别连接渐近平直时空。这同样也不是必需的，因为在非渐近平直时空中也可以有虫洞。但正如在非渐近平直时空中可以存在黑洞那样的东西，物理学家们研究黑洞时仍普遍假设时空是渐近平直的，虫洞研究也是如此。这一简化条件还可以这样来理解，那就是虫洞本身的结构与时空的大尺度结构并无密切关系，因此不妨对后者采用最便利的假设〔6〕。

为清楚起见，我们把上面提到的可穿越虫洞所需满足的所有条件罗列在一起〔7〕：

（1）［“经验教训”］不存在视界；

（2）［“经验教训”］穿越过程中遇到的应力是人体能够承受的；

（3）［一般条件］满足广义相对论场方程；

（4）［一般条件］物质的能量动量张量是物理上存在的；

（5）［一般条件］物质的数量是可观测宇宙可以提供的；

（6）［一般条件］在微扰下保持稳定；

（7）［简化条件］度规是静态球对称的；

（8）［简化条件］“喉咙”是唯一的；

（9）［简化条件］出入口分别连接渐近平直时空。

条件列出了，接下来就是寻找满足条件的具体虫洞解了。由于下面讨论的全都是可穿越虫洞，为行文简洁起见，有时将会略去“可穿越”这一限定词。在广义相对论中，寻找具体解的传统做法是首先给定物质分布（即物质能量动量张量的分布），然后求解广义相对论场方程以得到时空结构。这一做法体现的是物质为因、几何为果的物理思想，或者用惠勒的话说：“物质告诉时空如何弯曲，时空告诉物质如何运动。”不过对于虫洞来说，这种做法很不方便，因为虫洞的物质分布在索恩和莫里斯的研究之前乃是无人知晓的东西，倒是它的时空结构早在惠勒的概念性研究中就已经有了直观图示。因此，索恩和莫里斯采用了一个聪明的思路，那就是将传统做法逆转，即从时空结构入手，然后用广义相对论场方程计算出物质分布。这种逆转在数学上是完全等价的（颠过来倒过去都是广义相对论场方程），在物理上却有着微妙的差别，那就是传统做法由于首先给定了物质分布，因此可以直接保证物质分布是物理上可以实现的（即满足条件（4）和（5）），而逆转的做法却无法直接保证这一点。这一微妙差别导致的后果我们很快就会看到。

由于虫洞的出入口分别连接渐近平直时空（条件（9）），这启示我们引进两个坐标域（coordinate patch），分别描述出口和入口附近的时空，两者在“喉咙”处相互衔接。而度规因为是静态球对称的（条件（7）），其一般形式是众所周知的，包含两个任意函数，且两者都只是径向坐标r的函数。由于广义协变性，度规形式的选择有很大的自由度。对于我们来说，比较方便的做法是将之表述成与施瓦西度规有一定类似性的形式：

$$ds^2 = e^{2\varphi_\pm(r)}dt^2 - \left[1 - \frac{b_\pm(r)}{r}\right]^{-1}dr^2 - r^2 d\Omega^2 \qquad (5.3.1)$$

其中$\varphi_\pm$和$b_\pm$的下标±分别表示两个坐标域，坐标r的取值范围是［r_0，∞），r_0是“喉咙”所对应的径向坐标值。由于广义相对论场方程（条件

（3））是二阶微分方程，因此我们要求$\varphi_{\pm}$（r）和$b_{\pm}$（r）起码是二次可微的。我们并且还要求$\varphi_{\pm}$（r）处处有限，这是不存在视界（条件（1））的体现，因为它保证了$e^{2\varphi_{\pm}(r)}$不会像施瓦西度规那样在某些地方（即视界上）为零。

除上述一般限定外，这一度规在$r=r_0$和$r\to\infty$需要满足一些边界条件。对于$r\to\infty$，由于出入口连接渐近平直时空（条件（9）），因此φ_{+}（∞）和φ_{-}（∞）均为（有限）常数。在一般情况下，这两个常数可以是不相等的（请读者想一想，这两个常数不相等的物理意义是什么？）。同样的，b_{+}（∞）和b_{-}（∞）也均为（有限）常数。与施瓦西度规相对比不难看出，b_{+}（∞）和b_{-}（∞）分别对应于在出入口所连接的渐近平直时空中测得的出入口——也称为“嘴巴”（mouth）——的质量（确切地说是质量的两倍）。一般来说，这两个常数也可以是不相等的，即虫洞的两个“嘴巴”的质量可以是不相等的〔8〕。

对于$r=r_0$（即“喉咙”处），由于两个坐标域在此衔接，且$\varphi_{\pm}$（r）和$b_{\pm}$（r）起码是二次可微的。因此$\varphi_{+}(r_0)=\varphi_{-}(r_0)$，$\varphi'_{+}(r_0)=\varphi'_{-}(r_0)$，$b_{+}(r_0)=b_{-}(r_0)$，$b'_{+}(r_0)=b'_{-}(r_0)$。不仅如此，$r=r_0$作为虫洞的“喉咙”，是径向坐标取值最小的地方，因此此处沿径向的$dr/ds=[1-b_{\pm}(r)/r]^{1/2}=0$，而$d^2r/ds^2=[b_{\pm}(r)/r-b'_{\pm}(r)]/(2r)\geq 0$。这表明（请读者自行证明）

$$\begin{cases} b_{+}(r_0)=b_{-}(r_0)=r_0 \\ b'_{+}(r_0)=b'_{-}(r_0)\leqslant 1 \end{cases} \qquad (5.3.2)$$

由于“喉咙”是唯一的（条件（8）），因此在偏离但靠近“喉咙”的一个开区间$(r_0, r_0+\Delta)$内，$dr/ds>0$，$d^2r/ds^2>0$，或者等价地（也请读者自行证明）：

$$\begin{cases} b_{\pm}(r) < r \\ b'_{\pm}(r) < \dfrac{b_{\pm}(r)}{r} \end{cases} \qquad (5.3.3)$$

以上就是球对称可穿越虫洞的时空结构所需满足的一般条件。另一方面，度规的静态球对称（条件（7））也给物质能量动量张量的形式施加了一定的限制，使它在t、r及两个横向坐标组成的正交标架场中具有$T^{ab}=\mathrm{diag}(\rho, \tau, p, p)$的正则形式（请读者想一想，球对称体现在哪里？），其中ρ是能量密度，τ是径向张力，p是横向压强，它们都只是径向坐标r的函数。这是球对称可穿越虫洞的物质分布所需满足的一般条件。

下一步要做的就是所谓的“用广义相对论场方程计算出物质分布”，由于我们已将物质分布归结为ρ、τ、p这三个函数，因此这一步实质上就是用前面引进的球对称度规计算出相应的爱因斯坦张量，将之——依据广义相对论场方程（条件（3））——与$8\pi T^{ab}$等同起来，从而得到描述时空结构的函数φ、b与描述物质分布的函数ρ、τ、p之间的关系（这里我们丢弃了$\varphi_{\pm}$和$b_{\pm}$中表示坐标域的下标±，因为接下来的计算与坐标域无关）。这一计算是直截了当的，因为静态球对称度规的爱因斯坦张量的计算是广义相对论中的标准内容（当然，具体计算与所选择的度规形式有关），结果也并不复杂（感兴趣的读者可以自己推算一下）：

$$\begin{cases} 8\pi\rho = \dfrac{b'}{r^2} \\ 8\pi\tau = \dfrac{2(r-b)\varphi'}{r^2} - \dfrac{b}{r^3} \end{cases} \qquad (5.3.4)$$

其中比这两个方程更复杂的关于p的方程因后面不会用到而省略了。

由这两个方程可以得到两个重要结果。一个是在“喉咙”处，由上述第二个方程可知

$$\tau(r_0) = -\frac{1}{8\pi r_0^2} \tag{5.3.5}$$

这个结果对于探讨虫洞是否真的“可穿越”有着重要影响，我们将在后文中进一步讨论。另一个是将这两个方程相加：

$$8\pi(\rho+\tau) = \frac{b'}{r^2} + \frac{2(r-b)\varphi'}{r^2} - \frac{b}{r^3} = -\frac{e^{2\varphi}}{r}\left[e^{-2\varphi}\left(1-\frac{b}{r}\right)\right]' \tag{5.3.6}$$

由此不难看出在“喉咙”处（请读者自行证明）：

$$\rho(r_0) + \tau(r_0) = -\frac{[b'(r_0) - b(r_0)/r_0]}{8\pi r_0^2} \leqslant 0 \tag{5.3.7}$$

将这一结果与1.2节所介绍的能量条件相比较，可以看到零能量条件已是岌岌可危了——若不是“≤”中还包含有“＝”的话，就已经被破坏了。但这个“＝”管得了“喉咙”却管不了周围，因为式（5.3.6）中的$e^{-2\varphi}$（1－b/r）在“喉咙”处为零[9]，在偏离但靠近“喉咙”的一个开区间（r_0，r_0＋Δ）内却大于零（参阅式（5.3.3）），因此在偏离但靠近“喉咙”的一个开区间（r_0，r_0＋δ）内，［$e^{-2\varphi}$（1－b/r）］′＞0，从而

$$\rho(r) + \tau(r) < 0 \tag{5.3.8}$$

这表明在球对称可穿越虫洞的“喉咙”附近零能量条件会遭到破坏。由于零能量条件比弱能量条件、强能量条件、主能量条件都弱，因此它的破坏意味着在球对称可穿越虫洞的“喉咙”附近弱能量条件、强能量条件和主能量条件都会遭到破坏。这是索恩和莫里斯所得到的最重要的结果之一，也是我们在5.1节中提到的索恩在参加完毕业典礼返家途中得到的“喜忧参半的初步结果”中“忧”的部分。

我们前面提到过，索恩和莫里斯在寻找虫洞解的过程中对传统做法的逆转从数学上讲是等价的，在物理上却有着微妙差别，现在我们看到的正是这一微妙差别导致的后果。不过，这一后果虽有可“忧”之处，却也恰恰是索恩和莫里斯逆转传统做法的价值所在，否则的话——即首先

给定物质分布的话，是很难想到要引入违反能量条件的物质分布的，从而也就得不到球对称可穿越虫洞解了。至于违反能量条件是否意味着物质分布不再是物理上可以实现的（即违反条件（4）和（5）），我们在后文中将会进一步讨论。

关于上面这个结果，还有一个有趣的尾声可以提一下。索恩和莫里斯的论文提交后不久，索恩以前的学生、当时在美国宾夕法尼亚州立大学（Pennsylvania State University）物理系任教的佩奇（Don Page）给莫里斯写了一封信，提到索恩和莫里斯的上述结果可以很容易地从霍金和艾里斯（George Ellis）的名著《时空的大尺度结构》（The Large Scale Structure of Space-Time）的某些结果中推得。索恩和霍金是老朋友兼"赌友"（我们在4.1节末尾曾经介绍过他们打过的一个赌），彼此是很熟悉的，但索恩对霍金所擅长、并在那部名著中详加运用的所谓"全局方法"却并不熟悉。昔日学生佩奇——他同时也是霍金的学生——的这封来信使索恩既震动又惭愧，他在"检讨"这件事情时"沉痛"地表示："我感到多么愚蠢啊，我从未深入学习过全局方法（霍金和艾里斯书中的课题），现在我为此付出了代价。"

佩奇提到的方法确实要比索恩和莫里斯的容易许多。当然，这种"容易"是见仁见智的，因为那是"站在巨人的肩膀"（全局方法）上的容易。如果不熟悉全局方法，那它非但不容易，反而是相当困难的。不过，虽不曾强调过，我们在前面介绍奇点定理时所涉及的很多内容其实就属于全局方法。因此，我们恰好幸运地具备了"站在巨人的肩膀"上的便利。利用这种便利，让我们对佩奇提到的方法作一个简单介绍。这种方法的切入点是2.3节中的式（2.3.1）（即雷查德利方程）。这是描述测地线束的方程。对于我们的目的来说，测地线束中的测地线要选为类光测地线，固有时间τ要改为仿射参数λ，切矢量的记号V则要改为表示类光矢量的k，即

$$\frac{d\theta}{d\lambda} = -R_{ab}k^a k^b - \frac{1}{3}\theta^2 - \sigma_{ab}\sigma^{ab} \qquad (5.3.9)$$

可以证明，对于沿径向传播的类光测地线束，切变张量σ_{ab}恒为零，因此上式可以简化为$d\theta / d\lambda = -R_{ab}k^a k^b -$（1/3）$\theta^2$。由于膨胀标量θ描述的是测地线束的会聚或发散，因而在“喉咙”处，即测地线束从会聚转为发散的地方，θ＝0。另一方面，由于自喉咙向外测地线束是发散的，即θ＞0，因此在“喉咙”附近的一个小区域内$d\theta / d\lambda > 0$，从而$R_{ab}k^a k^b < 0$。而这等价于$T^{ab}k_a k_b < 0$（请读者自行证明），即破坏了零能量条件——与索恩和莫里斯的结果相同。

佩奇提到的全局方法不仅“容易”，而且还有一个更重要的优点，那就是不涉及度规的具体形式，从而可以被推广到更普遍——比如没有对称性——的情形。事实上，利用这类方法，物理学家们已经证明了远比索恩和莫里斯的结果普遍得多的结果，比如：全局双曲时空中的可穿越虫洞至少会在一条类光测地线上破坏零能量条件（从而也破坏弱能量条件、强能量条件和主能量条件）。由于全局双曲是一种非常优良、因而在一定程度上被认为是物理时空所具有的品质（虽然这种品质从理论上讲似乎太强了一点，而且极难有验证的可能）。因此这一结果表明可穿越虫洞对零能量条件（以及弱能量条件、强能量条件和主能量条件）的破坏很可能是普遍的〔10〕。

5.4　奇异物质——负能量的挑战

在上节中，我们得到了一个很可能适用于所有可穿越虫洞的普遍结论，那就是可穿越虫洞的物质分布（即物质能量动量张量的分布）会破坏零能量条件（从而也破坏弱能量条件、强能量条件和主能量条件）。物理学家们把那样的物质称为“奇异物质”（exotic　matter）[11]。很明显，奇异物质之所以“奇异”，就在于零能量条件遭到了破坏。但是，零能量条件的破坏本身到底又“奇异”在哪里呢？对普通读者来说恐怕不是一清二楚的，因为能量条件哪怕对于物理专业的人来说，也并不属于“日常用语”。在本节中，我们首先要做的，就是将零能量条件的破坏具体化，将它与物理专业乃至普通人的“日常用语”联系起来，以显示其“奇异”之所在。

我们知道（参阅1.2节），零能量条件要求对于任何一个主压强p，ρ＋p≥0。因此，零能量条件的破坏意味着起码对于某个主压强p，ρ＋p＜0。为了显示这一条件的物理实质，让我们引进一个沿p所代表的主方向以速度v运动的惯性系Σ。由能量动量张量T^{ab}＝diag（ρ，p）（这里我们略去了对当前计算没有影响的两个空间维度）的洛伦兹变换可以得到Σ系中的能量密度为（感兴趣的读者请自行计算一下）

$$\rho' = T^{00\prime} = T^{ab}\Lambda_a^0\Lambda_b^0 = \gamma^2(\rho + p) - p \qquad (5.4.1)$$

其中Λ为洛伦兹变换矩阵，γ＝$(1-v^2)^{-1/2}$是洛伦兹因子（Lorentz factor）。由于ρ＋p＜0，而γ可以任意大（只要速度足够接近光速），因此总可以选择惯性系Σ，使得ρ′＜0，即总可以选择惯性系Σ，使得其中观测到的能量密度小于零！这种负能量密度就是零能量条件的破坏所具有的“奇异”性质。由于这一性质，奇异物质也被称为负能量物质。

如果说得更浅白一点的话，那么奇异物质之所以“奇异”，是因为在

物理学中，能量的零点是用真空来定义的。相对于这一零点，任何其他状态（即有物质的状态）都具有正能量，物质越多，能量就越高；物质越少，能量则越低。依照这样的定义，负能量作为比零能量更低的能量，意味着比号称一无所有的真空具有“更少”的物质，这在经典物理学中是不可思议的，甚至近乎于语义上矛盾的说法。

假如这种“不可思议”或“语义上矛盾的说法”就是故事的终点，那么通过虫洞进行星际旅行或所谓“可穿越虫洞”的美丽设想也就可以算走到终点了。幸运的是，这种“不可思议”或“矛盾”来自于真空“一无所有”这样一个经典物理学的观念，而我们知道，经典物理学并不是物理学的终点，在它之后还有所谓的量子理论。量子理论的出现修正甚至颠覆了许多经典物理学的观念。那么，使负能量陷入困境的真空“一无所有”的观念是否也在此列呢？答案是肯定的，因为量子理论的发展彻底改变了我们对真空的理解。在量子理论中，真空所表示的乃是量子场的基态，它与其他状态（即有物质的状态）的区别，远没有像经典物理学中的“无”和“有”的区别那样截然。从某种意义上讲，量子理论中真空与其他状态的区别，不过是量子场能级的“低”和“高”之间的区别而已，定量多过定性。真空在量子理论中不仅像其他状态一样有自己的结构，而且还是高度动态的，随时可以有虚粒子对的产生和湮灭。在这种全新的、不再“一无所有”的真空下，负能量至少从概念层面上讲不再是不可思议，更不是语义上矛盾的了。

当然，这只是为负能量扫清了概念障碍，而并不足以确立它的实际存在。负能量的实际存在——如果可以确立的话——必须诉诸实验，最低限度——假如实验一时无法实现的话——也需要有更具体、更直接的理论支持。这是负能量对虫洞物理学的基本挑战。幸运的是，这种挑战得到了很正面的回应。经过理论物理学家与实验物理学家长期而共同的努力，理论支持与实验确立先后得以实现。

我们先来看看理论支持。既然量子理论的真空是有结构的，那么不

难想象，这种结构与几乎所有其他物理体系的结构一样，会受边界条件的影响。这就启示我们考虑特定边界条件下的真空，假如那种真空的能量比“普通”（即边界可以忽略的）真空具有更低的能量，那么就可以被视为一种负能量。沿这一思路的最著名的结果，是荷兰物理学家卡西米尔（Hendrik Casimir，1909—2000）提出，并以他名字命名的卡西米尔效应（Casimir effect）。1948年，卡西米尔发表了一篇题为“论两块理想导体板之间的吸引”（On the attraction between two perfectly conducting plates）的论文，对两块相互平行的理想导体板——以下将简称为平行导体板——之间的真空能量面密度（即单位面积导体板之间空间内的总真空能量）进行了计算。结果表明，那能量面密度确实是负的！更重要的是，那能量面密度与平行导体板的间距有关（这是可以预料的，因为当间距趋于无穷时，因边界条件而产生的能量面密度必须消失），从而意味着平行导体板之间会因这种能量的存在而产生一种相互作用力。这种相互作用力被称为卡西米尔力（Casimir force），它的存在为从实验上检验卡西米尔效应提供了直接途径。

平行导体板情形下卡西米尔效应的计算并不复杂，只不过是对有边界条件和没有边界条件下的量子场基态能量——即所谓的零点能（zero-point energy）——进行比较而已，并且已被一些标准量子场论教材所收录，就不在这里重复了[12]。不过，直接把结果罗列出来又显得太偷懒，因此我们决定采取“中庸之道”，介绍一种很有用的半定量方法：量纲分析法（dimensional analysis），用它来给出计算结果的主要部分。众所周知，理论物理所用的方法几乎清一色是数学方法，如果一定要从中找出一些具有“物理特色”的方法的话，那么量纲分析法或许就是其中之一。这一方法虽然也披着简单的数学外衣，但其核心是属于物理的，因为量纲是物理量特有的东西。量纲分析法主要由两个步骤组成：

（1）通过对物理意义的分析，列出结果有可能依赖的所有物理量和物理常数；

（2）通过对量纲的比较，确定结果对那些物理量和物理常数的具体依赖关系。

通过这两个步骤，运气好的话，可以将结果确定到只差一个常数（比例系数）的程度。平行导体板情形下的卡西米尔效应就是如此。

在卡西米尔效应的理论研究中，我们通常假定平行导体板足够大，从而导体板——作为二维客体——的边缘可以忽略〔13〕。在这种近似下，结果（即真空能量面密度）所能依赖的物理量只有一个，那就是平行导体板的间距，我们用d来表示，其余的就都是物理常数了。由于导体板所施加的边界条件是针对电磁场的，因此卡西米尔效应所涉及的量子场是电磁场，零点能则是电磁场的零点能，这种零点能只依赖于两个物理常数：光速c和普朗克常数ħ〔14〕。因此，平行导体板情形下卡西米尔效应中的真空能量面密度U只依赖于平行导体板的间距d、光速c和普朗克常数ħ，即U＝f（d，c，ħ）。这就完成了两个步骤中的第一个。

那么具体的依赖关系是什么呢？这就需要对量纲进行比较了。为了简洁起见，我们用［X］表示物理量或物理常数X的量纲，用M、L和T分别表示质量、长度和时间的量纲，则

$$\begin{cases} [U] = \mathrm{MT}^{-2} \\ [d] = \mathrm{L} \\ [c] = \mathrm{LT}^{-1} \\ [\hbar] = \mathrm{ML}^2\mathrm{T}^{-1} \end{cases} \qquad (5.4.2)$$

很明显，唯一能让函数关系U＝f（d，c，ħ）的左右两边具有相同量纲的函数形式是（请读者自行证明）

$$U \propto \frac{\hbar c}{d^3} \qquad (5.4.3)$$

这就完成了第二个步骤。可惜，用量纲分析法无法得到比例系数（这是该方法被称为“半定量方法”的原因所在）。对零点能的具体计算表明，

比例系数为-π^2 / 720。这样，我们就得到了真空能量面密度的确切表示式，即

$$U = -\frac{\pi^2 \hbar c}{720 d^3} \qquad (5.4.4)$$

相应的体密度和能量动量张量则为

$$\rho = \frac{U}{d} = -\frac{\pi^2 \hbar c}{720 d^4} \qquad (5.4.5)$$

和（假定平行导体板的法向为z方向）[15]

$$T^{ab} = -\frac{\pi^2 \hbar c}{720 d^4}\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 3 \end{pmatrix} \qquad (5.4.6)$$

由能量面密度，立刻可以得到平行导体板之间沿法向的相互作用力的面密度（即作用于单位面积导体板上的力）为

$$F = -\frac{\partial U}{\partial d} = -\frac{\pi^2 \hbar c}{240 d^4} \qquad (5.4.7)$$

其中结果中的负号表示这种力是相互吸引的。

以上就是有关平行导体板情形下卡西米尔效应的主要结果。在卡西米尔效应的研究中，有一段小小的历史值得一提。比针对平行导体板情形的研究稍早，卡西米尔与荷兰物理学家玻德（Dirk Polder，1919—2001）曾合作发表过一篇论文，以类似于推导分子间范德瓦尔斯力（van der Waals force）的方法，在大距离极限下，对理想导体板与单个的原子或分子，以及原子或分子彼此之间的相互作用进行了研究。那项研究通常被视为研究卡西米尔效应的先导，它所采用的方法也被视为对卡西米尔效应的另一种理解[16]。在那项先导研究的进行期间，卡西

米尔有一次访问哥本哈根（Copenhagen）时遇到了著名丹麦物理学家玻尔（Niels Bohr，1885—1962）。玻尔问了一句物理学家们见面时的“标准问候”：你最近在研究什么？卡西米尔就将他与玻德正在进行的研究告诉了玻尔，并表示自己希望能用更简单、更优美的方法来进行推导。玻尔思考了一下，然后——依照卡西米尔的回忆——嘟哝了一句：“肯定跟零点能有点关系。”姜不愧是老的辣！这句“嘟哝语”给了卡西米尔极大的启示，使他很快通过对零点能的计算，完成了平行导体板这一新情形的独立研究（即得到了我们上面介绍的若干结果）。这一研究因避免了原先依赖的大距离极限，而——如我们很快将会看到的——大大增加了实验验证的可能性。

卡西米尔效应是一种非常微小的量子效应（这并不奇怪，量子效应本就是以微小著称的）。从上面的公式中可以很容易地计算出有关物理量的数值（当然，一切都是跟平行导体板的间距有关的）。比如能量密度ρ——由式（5.4.5）可知——约为

$$\rho = -\frac{4.3\times10^{-28}}{d^4} \qquad (5.4.8)$$

这里所有的物理量都采用了国际单位制（SI）中的单位，即能量密度ρ以焦耳每立方米（J/m^3）为单位，平行导体板的间距d以米（m）为单位。如果用质能关系式将能量密度折合成质量密度，则约为

$$\rho = -\frac{4.8\times10^{-45}}{d^4} \qquad (5.4.9)$$

其中质量密度ρ以千克每立方米（kg/m^3）为单位。这无疑是非常微小的密度。类似地，卡西米尔力的面密度也非常微弱，由式（5.4.7）可知，约为

$$F = -\frac{1.3\times10^{-27}}{d^4} \qquad (5.4.10)$$

其中面密度F以牛顿每平方米（N/m^2）为单位。

接下来简单提一下卡西米尔效应的实验确立。前面说过，卡西米尔力的存在为从实验上检验卡西米尔效应提供了直接途径。但是途径虽然有了，测量由式（5.4.10）所表示的那么微小的力对实验物理学家来说依然是极大的挑战。唯一有利的，是这种力反比于平行导体板间距的四次方，从而可以通过缩小平行导体板的间距而得到非常显著的“放大”，这堪称实验验证中一个最重要的努力方向（前面提到的因避免大距离极限而大大增加实验验证的可能性，原因正在于此）。比如平行导体板的间距若能缩小到微米（μm）量级，力密度就可以增加到千分之一牛顿每平方米的量级。只不过，让两块导体板靠得如此之近而又互不接触绝非易事，通常只有对很小的导体板才能做到。而导体板越小，作用在导体板上的力的总量也将变得越小，从而又反过来增加了检测难度。不过幸运的是，卡西米尔效应并不局限于平行导体板，而早已被理论物理学家们推广到了各种更复杂的情形，比如针对其他形状、非理想导体板、电介质板、有限温度，等等，为实验研究提供了多种选择。那些更复杂的情形由于偏离理想条件而使得理论计算更为复杂，但对于实验验证来说，却因为所要求的条件不那么苛刻，而变得更容易实现。唯一不变的是，无论哪种情形下的卡西米尔效应都是非常微小的。因此，对卡西米尔效应的实验研究虽然从20世纪50年代起就陆续有人在做，却直到90年代后期才开始有了较高精度的验证〔17〕。

除卡西米尔效应外，还有一些其他量子效应也能在某些特定区域产生负能量。因此，奇异物质的存在是毋庸置疑的，这对于可穿越虫洞来说无疑是“利好消息”。只不过，这个“利好消息”是否真能让我们“获利”，却是大为可疑的。因为所有那些（源于量子效应的）负能量都是非常微小的，而且全都依赖于非常特殊的环境条件的配置才会出现，从而完全不像普通物质那样可以被随心所欲地、独立于其他物体地移动，并建造我们所需要的结构。从这个意义上讲，奇异物质的“奇异”虽是货

真价实的，能否对得住“物质”这一“光荣称号”却大可商榷。起码，它远没有普通物质那样的独立性。奇异物质的存在离不开以特殊方式配置起来的环境条件，它的移动也必须伴随环境条件的变更，就像垂危病人的移动必须伴随着医疗器械的移动一样。而那些无可或缺的环境条件——它们都是由普通物质构成的——完全有可能对奇异物质的“效用”（主要是引力效应）产生重大干扰，甚至彻底地抵消和淹没之〔18〕。因此，负能量（或奇异物质——如果我们依然愿意这么称呼它的话）虽然确实是存在的，它能否被用来构筑可穿越虫洞，却是一个很大的未知数——而且恐怕是极不容乐观的未知数。

5.5　虫洞的“工程学”

在5.4节中，我们将可穿越虫洞的物质分布会破坏零能量条件这一结论具体化了，那就是构筑可穿越虫洞需要用到所谓的奇异物质（或负能量物质）。并且以卡西米尔效应为例，对奇异物质的产生方式做了介绍。我们看到，奇异物质由于是靠量子效应产生的，其数量往往是非常微小的。卡西米尔效应从理论上被提出到实验上被验证，相隔了差不多半个世纪的时间，主要原因也正在于此。这种数量上的微小对于构筑可穿越虫洞来说当然是坏消息，但究竟坏到什么程度呢？还得看“供求关系”的另一边——即构筑可穿越虫洞究竟需要多少奇异物质。如果需要量很少，情况也许就不算太坏；反之，则不容乐观。

那么，构筑可穿越虫洞究竟需要多少奇异物质呢？在本节中，我们将转入虫洞研究的所谓“工程学”方面，来对这一问题作些探讨。

很明显，构筑不同结构的虫洞所需要的奇异物质的数量可能是不同的，因此这一问题的答案取决于虫洞的具体结构。不过，涉及虫洞具体结构的计算通常是不太容易的，而且那种不太容易的计算往往还没什么特别的重要性（因为往往没什么理由认为某种结构比另一种结构更值得计算），从而给人一种吃力不讨好的感觉。怎么办呢？物理学家们想到了一个聪明的点子，那就是近似——既然选不出最值得计算的虫洞结构，不如只做近似计算，起码还能图个简单。在广义相对论研究中，有一种很常用的近似叫做“薄层”（thin　shell）近似，它的要点是将物质分布近似为薄层分布，虫洞物理学上的许多计算就采用了这一近似。在数学上，所谓的薄层是被假定为无限薄的（即厚度趋于零），相应的物质分布呈现δ函数的形式。

薄层近似不仅具有数学上的简单性，从“工程学”的角度讲也不无优点。因为尽管如我们在5.4节末尾所说，奇异物质能否被用来构筑可穿

越虫洞是一个“极不容乐观的未知数”，但假如它能够，那么从“工程学”的角度讲，很自然的做法就是将这种数量非常微小的奇异物质的用量尽可能减少。而减少用量的途径之一就是将它的使用范围尽可能局限起来，比如局限在薄层上。除此之外，薄层近似还有一个很大的优点，那就是它在数学上的简单性使人们可以在其他方面适当地引进复杂性，比如引进非球对称性。我们在后文中将会看到，这对于更全面地分析虫洞的可行性是有很大益处的。

利用薄层近似，物理学家们对构筑可穿越虫洞所需要的奇异物质的数量进行了计算。计算的结果虽仍与虫洞的具体结构——即薄层形状——有关，但托薄层近似的福，可以用一个系统的公式来概括了。在这里，我们将再次采取“中庸之道”，用量纲分析法来推出那个系统的公式。我们用M表示构筑可穿越虫洞所需要的奇异物质的数量。很明显，M与虫洞喉咙的大小——也称为虫洞的半径——r_0有关。除此之外，作为广义相对论的结果，它当然也跟万有引力常数G及光速c有关，因此M＝f（r_0，G，c）。与上节中推导卡西米尔效应的做法完全类似，我们可以通过对量纲进行比较，得到唯一的函数形式（感兴趣的读者不妨自己推导一下），即

$$M \propto -\frac{c^2}{G} r_0 \qquad (5.5.1)$$

这里，我们将负号单独列出，以表示奇异物质的能量为负这一重要特征。考虑到——如前所述——计算结果与薄层形状有关，因此完整的结果还需包含一个与薄层形状有关的因子F_s，即

$$M = -\frac{c^2}{G} F_s r_0 \qquad (5.5.2)$$

其中比例系数已被吸收进了F_s之中，因此比例符号“∝”变成了等号“＝”。F_s作为与可穿越虫洞的具体结构有关的因子，承载了计算中的

细节，但在数量级意义上对结果没有显著影响。式（5.5.2）对于更一般的虫洞也基本成立，只不过F_s的计算通常要复杂得多。

将万有引力常数G及光速c的数值代入，不难得到与上述公式相对应的数量级意义上的数值结果（感兴趣的读者可以核验一下）

$$M = -M_{\odot} r_0 \qquad (5.5.3)$$

其中$M_{\odot}$为太阳的质量（数量级为10^{30}kg），虫洞的半径r_0以千米（km）为单位。

把这一结果与上节所得到的有关奇异物质质量密度的数值结果式（5.4.9）相比较，我们看到，构筑可穿越虫洞在“工程学”上绝对是一个空前巨大的挑战。因为太阳作为一个天体，它的质量乃是所谓的“天文数字”，而奇异物质来自量子效应，其数量是典型的“微乎其微”。因此，构筑可穿越虫洞乃是要将“微乎其微”的量子效应累积成“天文数字”般的巨大数量，其难度是可想而知的（更不用说这种量子效应能否被累积还是一个未知数）。

更糟糕的是，难度上的挑战并未到此为止。因为依靠相当于一个太阳质量的奇异物质所能构筑的，乃是一个半径约1000米的虫洞。而接下来的问题是：那样的虫洞真的可穿越吗？

初看起来，1000米算是不小的半径，假如虫洞类似于普通隧道的话，那样的半径应该足以让大小可观的星际飞船穿越了。看过科幻影片的人想必都对星际飞船穿越虫洞的特技处理留有深刻印象。从影片中看，星际飞船的周围充斥着星光组成的绚丽幻象，它所穿越的则似乎只是一条狭小的通道（见图5.2）。但实际上，星际飞船穿越虫洞的情形远比科幻影片所设想的来得复杂。为了让飞船及乘员安全穿越虫洞，几何半径的大小并不是最主要的问题。那么什么才是最主要的问题呢？是“活着”——也就是5.3节列出的可穿越虫洞所需满足的条件之第（2）条：穿越过程中遇到的应力是人体能够承受的。



图**5.2** 星际飞船穿越虫洞的特技处理

下面我们就以球对称虫洞为例，来谈谈这一条件。具体地说，穿越虫洞时，星际飞船及乘员将会遇到两种不同类型的应力：一种来自虫洞物质本身的张力，另一种则是虫洞引力场所产生的潮汐力。

我们先看前者。在5.3节中，我们曾经得到过一个当时称之为“对于探讨虫洞是否真的‘可穿越’有着重要影响”的公式，即式（5.3.5），它给出的是球对称可穿越虫洞喉咙处的物质张力。考虑到虫洞解的连续性，该公式也近似地给出了喉咙附近一个小区域内的物质张力。由于球对称可穿越虫洞的喉咙附近是零能量条件遭到破坏的区域（参阅5.3节），从而也就是奇异物质的分布区域（参阅5.4节），因此该公式给出的张力也被称为是奇异物质的张力〔19〕。

我们为什么将这个张力称为“对于探讨虫洞是否真的‘可穿越’有着重要影响”呢？关键在于它的数值。这数值从式（5.3.5）中是难以直接看出的，因为物理常数被略去了。现在我们就将之恢复起来。恢复的方法大家应该已经很熟悉了，就是量纲分析法。恢复的结果是（请读者自行验证）

$$\tau = \frac{c^4}{8\pi G r_0^2} \qquad (5.5.4)$$

这里我们略去了等式左端表示喉咙位置的r_0，以及等式右端来自约定，

从而并无本质意义的负号。将物理常数的数值代入便可得到数值结果为

$$\tau = \frac{5 \times 10^{10}}{r_0^2} \qquad (5.5.5)$$

其中张力τ以牛顿每平方米（N/m^2），或等价地以焦耳每三次方米（J/m^3）为单位，虫洞半径r_0以光年（light year）为单位。

这里需要特别注意的是，虫洞半径所采用的不是像“米”或“千米”之类的日常单位，而是光年这样的巨大单位。因此，上式表明一个半径为一光年的球对称虫洞喉咙附近的张力大小约为$5\times10^{10}N/m^2$。那么$5\times10^{10}N/m^2$又是一个什么概念呢？它相当于在每平方米的面积上压上500万吨的重物，或大致相当于物质在原子线度上所能承受的张力〔20〕。这意味着半径为一光年的球对称虫洞喉咙附近的张力如果作用在物质上，将足以破坏原子结构〔21〕。由于物质的宏观结构——以所能承受的张力而论——要比原子结构脆弱得多，因此，无论星际飞船还是乘员，在那样的张力作用下都将是“不堪一击”的。而倘若虫洞的半径不是一光年，而是更小，则喉咙附近的张力将会更大（因为张力的大小与虫洞半径的平方成反比）。倘若虫洞的半径只有1000米，则张力将高达不可思议的$5\times10^{36}N/m^2$，相当于每平方米的面积上压上5亿亿亿亿吨的重物，这无论从“工程学”的意义上，还是从已知物质的性质上讲都是近乎荒谬的。

由此我们看到，别说是半径1000米的球对称虫洞，就连半径一光年的球对称虫洞喉咙附近的张力如果作用在物质上，也并非星际飞船或人体所能承受的。那我们还能依靠什么呢？就只能靠式（5.5.5）所给出的张力与虫洞半径的平方成反比这一特点，通过半径比一光年更大的虫洞来进行星际旅行了。但这在“工程学”上同样是近乎荒谬的。因为式（5.5.3）表明构筑可穿越虫洞所需的奇异物质的数量正比于虫洞半径，而且构筑半径为1000米的虫洞所需的奇异物质的数量就相当于一个太阳

的质量。由于一光年约为十万亿千米，因此构筑一个半径为一光年的虫洞所需的奇异物质的数量约相当于太阳质量的十万亿倍！构筑能让星际飞船及乘员平安穿越的真正意义上的球对称可穿越虫洞所需的奇异物质的数量甚至比这更多。如此数量的物质（比整个银河系的质量还大得多）别说是奇异物质，就算是普通物质也实在是惊世骇俗的。

在索恩应萨根的电话咨询而研究虫洞物理学的时候，他曾经用一个特殊的视角来判断虫洞作为星际旅行通道的可行性，那就是：什么样的事情是物理定律允许无限发达的文明（infinitely advanced civilization）去做的？之所以采用这样的视角，是因为作为理论物理学家，他所关注的仅仅是物理定律是否允许可穿越虫洞存在，而并不在意具体的实施难度——他把后者扔给了“无限发达的文明”去操心，因为“无限发达的文明”按定义就是能做到物理定律所许可的一切事情的。他并且把由这一特殊视角所提出的问题统称为“萨根式问题”（Sagan-type question）〔22〕。不过，即便采用“萨根式问题”的特殊视角，上面提到的构筑球对称可穿越虫洞所需要的奇异物质的数量仍是一个灰色地带，甚至可以说是对“萨根式问题”本身的一种诘难。因为单纯从物理定律上讲，那样的可穿越虫洞或许是可以存在的，但实际上却几乎可以断定，操控数量如此之巨的奇异物质对于生活在可观测宇宙中的任何文明——哪怕是“无限发达的文明”——都是不太可能的。因此，上述可穿越虫洞可以说是介于理论上不可能与实际上不可能之间。而萨根所设想的“纵横交错的虫洞”用上述可穿越虫洞来衡量，则是天方夜谭。

接下来再看看星际飞船及乘员将会遇到另一种应力：虫洞引力场所产生的潮汐力。潮汐力的根源是引力场的不均匀性。由于这种不均匀性，物体上两个相距Δx的点所感受到的引力加速度会有所不同。我们知道，引力加速度本身的效应是可以通过等效原理消去的，但这种不同点所感受到的引力加速度a的不同（即Δa）是无法消去的，从而将体现为一种额外的作用，那就是潮汐力。潮汐力是一种常见的力，因为它在地

球上就有一种明显的表现形式：潮汐（其名称也由此而来）。潮汐力的计算是广义相对论中的标准内容，其加速度的表达式为

$$(\Delta a)^{\mu} = -R^{\mu}_{\alpha\nu\beta}V^{\alpha}(\Delta x)^{\nu}V^{\beta} \qquad (5.5.6)$$

其中R是引力场的曲率张量，V是物体的四维速度。对于球对称虫洞，上述加速度可以分解为径向部分（Δa）$_{\parallel}$与横向部分（Δa）$_{\perp}$，结果分别为（感兴趣的读者可以用度规式（5.3.1）具体算一下）

$$(\Delta a)_{\parallel} = \left\{\left(1-\frac{b}{r}\right)\left[-\varphi''-(\varphi')^2\right]+\frac{(b'r-b)\varphi'}{2r^2}\right\}(\Delta x)_{\parallel} \qquad (5.5.7)$$

$$(\Delta a)_{\perp} = \frac{\gamma^2}{r^2}\left[(r-b)\varphi'+\frac{v^2}{2}\left(b'-\frac{b}{r}\right)\right](\Delta x)_{\perp} \qquad (5.5.8)$$

其中v是物体的三维速度值，γ＝（1－v^2）$^{-1/2}$是洛伦兹因子。

有了这些结果，原则上就可以计算出星际飞船及乘员穿越球对称虫洞时将会遇到的潮汐力了。但是与计算奇异物质数量的情形相似，结果显然也是跟球对称虫洞的具体结构有关的（此外还可能跟星际飞船的运动方式有关——因为跟飞船运动有关的v和γ也出现在了公式中）。我们采用的点子也类似，那就是近似。不过，原先用过的薄层近似在这里是不恰当的，因为它所具有的δ函数形式的物质分布反映到潮汐力上会产生奇异性。那么，什么样的近似适合潮汐力的计算呢？最简单的一种是所谓的近施瓦西（proximal Schwarzschild）近似。这是对施瓦西度规的一个貌似细微的变更——即变更为

$$ds^2 = \left(1-\frac{2m}{r}+\frac{\varepsilon}{r^2}\right)dt^2-\left(1-\frac{2m}{r}\right)^{-1}dr^2-r^2d\Omega^2 \qquad (5.5.9)$$

其中ε＞0是一个很小的参数[23]。很明显，除了r＝2M附近的一个小区域外，近施瓦西度规处处都很接近施瓦西度规。不过，ε从数值上讲虽然细微（即所谓“貌似细微”），对时空的整体性质却有着非同小可的影

响，因为引进了ε之后，r＝2M处的g_{tt}就不再为零，从而使得原本出现在r＝2M处的视界不复存在了。这就克服了将爱因斯坦-罗森桥视为星际旅行通道所遭遇的本质上是来自视界的困难。我们在5.2节中曾经说过，爱因斯坦-罗森桥所具有的连接两个渐近平直时空的特性，起码在表观上与虫洞有着异曲同工之处。来自视界的困难一旦被克服，这种“异曲同工”性就可以由表观转变为实质，爱因斯坦-罗森桥也就荣升为虫洞了。这种虫洞被称为近施瓦西虫洞，它的喉咙位于r＝2M处〔24〕。

由于近施瓦西虫洞的度规在远离喉咙处与施瓦西度规十分接近，因此潮汐力也就基本上等于施瓦西度规中的潮汐力，其加速度的表达式为（这也是广义相对论中的标准结果，感兴趣的读者可通过式（5.5.7）和式（5.5.8）进行验证）

$$(\Delta a)_{\parallel} = \frac{2GM}{r^3}(\Delta x)_{\parallel} \tag{5.5.10}$$

$$(\Delta a)_{\perp} = -\frac{GM}{r^3}(\Delta x)_{\perp} \tag{5.5.11}$$

其中横向加速度$(\Delta a)_{\perp}$的表达式带有一个负号，表示物体在横向上会被压缩；而径向加速度$(\Delta a)_{\parallel}$的表达式不带负号，表示物体在径向上会被拉伸。我们前面提到过，半径为1000米的虫洞喉咙附近的张力会达到不可思议的数值，从而无论从“工程学”的意义上，还是从已知物质的性质上讲都是近乎荒谬的。现在我们可以以近施瓦西虫洞为例，从潮汐力的角度来检视一下半径为1000米的虫洞。因为潮汐力的作用，当星际飞船接近虫洞时，飞船上的乘员会渐渐感觉到自己的身体在沿虫洞的方向（即径向）上被拉伸，而在与之垂直的方向（即横向）上被挤压。这种感觉在一开始只是让人稍有些不适而已。但由于潮汐力的大小反比于距离的三次方，因此随着飞船的继续前行，潮汐力会迅速增加，距离每缩小到十分之一，潮汐力就会增加至1000倍。当飞船距离虫洞中心还有

1000千米的时候，潮汐力的大小就基本达到了人体所能承受的极限。如果这时候飞船还不立刻折回的话，所有乘员都将在致命的潮汐力作用下丧命。再往前飞行一段距离，则飞船本身也将在潮汐力的作用下解体。最终从虫洞另一端飞出的将不是意气风发的星际旅行家，而是飞船和人体的残骸[25]！

这就是试图穿越半径1000米的近施瓦西虫洞的星际旅行家的下场。因此，从潮汐力的角度看，半径1000米的虫洞——起码就近施瓦西虫洞这一例子而言——是无法作为星际旅行的通道的。不仅如此，包含潮汐效应在内的虫洞的巨大引力场会在周围一个很大的范围内产生毁灭性的破坏力，科幻小说或电影中的那些出入口位于行星表面的虫洞从这个意义上讲，都是极不可能的。在萨根的故事中，曾有人反对将外星智慧生物传来的蓝图付诸实施，因为他们担心那可能是一个能够毁灭地球的装置，他们的担心是很有道理的。那么，什么样的虫洞产生的潮汐力才是在整个穿越过程中都是人体能够承受的呢？对近施瓦西虫洞来说，大约要半径在几万千米以上才行（读者可以自行估算一下）。这与来自张力的条件（半径一光年以上）相比虽然宽松得多，依然是非常大的尺度（相应地，所需要的奇异物质的数量则相当于太阳质量的几万倍以上）。

以上就是对虫洞的所谓“工程学”讨论，就讨论中所涉及的例子而言，基本结论是：可穿越虫洞的半径必须很大，才能保障穿越过程中遇到的应力是人体能够承受的。相应的，所需奇异物质的数量也必须很大——大到几乎没有实现的可能。不过，这一结论并非无懈可击，因为讨论中涉及的例子毕竟是有限的。那么，在更一般的情形下是否有可能绕过这一结论呢？对此我们在后文中还会做进一步讨论。

5.6　由虫洞到时间机器

关于虫洞物理学，还有一个十分有趣的衍生课题，是我们要在本节中讨论的，那就是虫洞与时间机器的关系。这种关系听起来有些玄妙，对于熟悉相对论的人来说其实是不难想到的，因为虫洞作为连接两个时空区域的捷径，可以说是在某种意义上实现了超光速旅行。或者换一个说法，是在某种意义上使得类空事件之间的物理联系（因果联系）成为了可能〔26〕。另一方面，众所周知的是：在相对论中，类空事件的时序是与参照系的选择有关的。因此虫洞从某种意义上讲可以说是开启了因果颠倒的大门，而这正是时间机器的重要特征。因此，虫洞与时间机器是有可能存在联系的。

不过，在索恩和莫里斯研究虫洞时，也许是因为对虫洞本身倾注了太多注意力的缘故，忽视了像虫洞与时间机器的关系那样的衍生课题。但幸运的是，那样的忽视并没有持续太久。1986年底，索恩和莫里斯在芝加哥参加了一次天体物理学会议，身为老板的索恩做了一个学术报告（但不是有关虫洞的），当时还是研究生的莫里斯则四处与人聊天，混个脸熟，其中有位聊天对象是中康涅狄格州立大学（Central Connecticut State University）的物理学家，名叫罗曼（Thomas Roman）。索恩和莫里斯有关虫洞的研究当时已接近完成，莫里斯就在聊天中介绍了该项研究。正所谓“当局者迷，旁观者清”，罗曼一听完莫里斯的介绍马上就问道：如果虫洞能让人以比光速更快的速度跨越空间距离，那是否意味着能用它旅行到过去？“一语点醒梦中人”，罗曼的问题让索恩和莫里斯大为震动，以至于索恩后来述及此事时再次“沉痛”检讨了自己的愚蠢〔27〕。

罗曼的问题使索恩和莫里斯立刻投入到了对虫洞与时间机器关系的研究之中（值得一提的是：学术交流是互惠的，受聊天的影响，罗曼自

己后来也涉足了虫洞与时间机器领域的研究），并在论文中添加了一个新的部分，叙述了一种利用两个虫洞构筑时间机器的方案。不久之后，他们又提出了一个新方案，将两个虫洞减少为一个。那个新方案就是目前文献中用虫洞构筑时间机器的标准方案（当然，不同文献在具体表述上各有特色，但基本思路是一致的）。

下面我们就对这个标准方案作一个简单介绍。这个方案是通过三个步骤构筑时间机器的：

（1）构筑一个虫洞（或利用一个现成的虫洞）。在介绍时间机器的构筑时，为简单起见，人们通常略去虫洞本身的“工程学”细节，而将之抽象为一个对相距L的两条世界线（t，0，0，0）和（t，0，0，L）作等时连接的通道〔28〕。

（2）设法在虫洞的入口和出口之间产生一个时间差。这有许多办法可以做到。比如可以通过让出口相对于入口作高速运动后返回，利用狭义相对论的时钟延缓效应来做到〔29〕，或通过将出口置于与入口不同的外部引力场中，利用引力场的时钟延缓效应来做到，等等。假定所产生的时间差为T，则这一步完成时虫洞所连接的两条世界线将成为（t，0，0，0）和（t＋T，0，0，L）。

（3）将虫洞出入口之间的距离L缓慢地缩小，直至接近于零〔30〕。这一步完成后，虫洞所连接的两条世界线将变成（t，0，0，0）和（t＋T，0，0，0）。

由上述三个步骤所构筑的连接（t，0，0，0）和（t＋T，0，0，0）的虫洞就是一种时间机器，因为由入口往出口方向穿越这一虫洞会跨越幅度为T的时间，由出口往入口方向穿越这一虫洞则会跨越幅度为-T的时间。因此，这一时间机器既可以通向未来，也可以重返过去〔31〕。

用虫洞构筑时间机器对于科幻作家来说可谓是喜出望外的事情，因为时间机器无论对于科幻作家还是他们的读者，都是吸引力堪与虫洞相比——甚至犹有过之——的东西（谁不想遨游未来或重返过去呢？）。

事实上，早在虫洞、黑洞甚至相对论问世之前，科幻作家们就已创作出了不少有关时间机器或时间旅行的故事。比如著名英国科幻作家威尔斯（H. G. Wells，1866—1946）早在1895年（即比狭义相对论的问世还早了十年）就发表了名为《时间机器》（The Time Machine）的著名小说。用虫洞构筑时间机器的设想虽问世很晚，但由于是从科学土壤里萌发出来的，且与虫洞这一热门概念紧密相通，因此大有后来居上的势头。在新的科幻小说或影片中，像普通机器那样“老土”的时间机器已基本淡出，取而代之的往往是用类似于穿越虫洞的文字或镜头表示的时间旅行〔32〕。

但物理学家们的态度要谨慎得多，因为时间机器在物理学上一向是烫手的山芋。其中最烫手的就是由此引发的几乎无穷无尽的因果悖论〔33〕。那些悖论是如此棘手，以至于在很长的时间里，时间机器被多数物理学家视为是幻想领域的东西，而非严肃课题。比如索恩研究时间机器的消息传出后，就曾接到过朋友的电话垂询，担心他会“走火入魔”。甚至他本人也担心虽然自己“老了，无所谓”，但此类消息有可能会影响学生的前程。当然，经过若干位物理学家的共同努力，如今的情况已有所改变，但时间机器依然是一个几乎在所有方面都有着高度争议性的话题。索恩自己的态度也非常谨慎。我们知道，索恩曾经跟霍金就宇宙监督假设打过赌，其中霍金的立场是“上帝憎恶裸奇点”。霍金对时间机器有一个类似的立场，叫做“自然憎恶时间机器”（Nature abhors a time machine）。用学术术语来说，霍金的立场叫做“时序保护假设”（我们在2.6节末尾提到过这一假设及其依据），它通俗地说就是认为时间机器起码在宏观尺度上是不存在的。对于霍金的这一立场，索恩也曾考虑过打赌，但思虑再三后，却未敢下注。其原因用他自己的话说，是“我喜欢跟霍金打赌，但只打我有合理可能性会赢的赌”。索恩并且承认，就时序保护假设而言，无论从直觉还是亲自做过的计算来看，他都觉得“霍金有可能是对的”。

如果霍金是对的，即“自然憎恶时间机器”，那么用虫洞构筑时间机器就应该是不可能的。但索恩——如上面三个步骤所显示的——又明明已经至少在理论上找到了构筑时间机器的方法，这该做何解释呢？对此，美国物理学家沃尔德和杰罗奇提出了一种可能性，那就是虫洞一旦被构筑成时间机器，会在构筑成功的一瞬间就自动毁灭。因为在那一瞬间，任何细微的辐射都可以通过虫洞返回过去，与它自身相叠加。这种叠加过程可以在零时间内重复无穷多次，从而产生毁灭性的自激效应〔34〕。

除去理论层面的考虑外，用虫洞构筑时间机器还面临着实际层面上的诸多困难。事实上，构筑可穿越虫洞本身就已如此困难，通过种种操作在其出入口之间产生可观的时间差无疑更是难上加难。从这个角度看，用虫洞构筑时间机器的前景是相当渺茫的。

5.7 讨论

我们有关虫洞物理学的介绍到这里就大体结束了。依照前面各专题所遵循的结构惯例，这最后一节将做一些略带延伸性的讨论。

首先要讨论的话题是：虫洞究竟有没有可能被构筑出来？读者也许会觉得奇怪？我们不是一直就在讨论这个问题，并且还得出了诸如构筑可穿越虫洞需要奇异物质之类的具体结论吗？怎么又绕回去了呢？不是绕回去，而是不同的问题。因为我们前面讨论的其实只是如何维持一个虫洞，而不是如何构筑一个虫洞，后者是指从无到有地产生一个虫洞，因而是完全不同的问题。

回到5.1节用过的苹果的比喻上来：如果人类生活在苹果的表面上，而虫洞是穿越苹果内部的通道，那么很明显，那样的通道是对苹果表面拓扑结构的一种改变。因此，从无到有地产生虫洞意味着“拓扑改变”（topology change）。关于这种拓扑改变的含义，物理学家们曾有过分歧。虫洞概念的“开山祖师”惠勒思考这一概念时，所针对的乃是时空，即认为时空的拓扑结构有可能因量子涨落而改变（参阅5.2节）。这个想法曾被一些物理学家所继承，但也有一些物理学家表示了怀疑，因为稍稍细想一下，就会发现这个想法是有问题的。所谓时空，顾名思义，乃是时间与空间的合称，是已经包含了时间的。既然如此，时空作为一个四维流形，按定义就是完全确定的，它的任何性质都不可能发生变化——因为在它之外并不存在可用来定义“变化”的时间。

那么，拓扑改变指的究竟是什么呢？比较合理的看法是将时空流形分解为空间部分和时间部分，然后将拓扑改变定义为空间部分的拓扑相对于时间的变化[35]。

那么，如此定义的拓扑改变有可能出现吗？物理学家们已经进行了一些研究，并得到了一些结果。比如1967年，杰罗奇（又是此人！）证

明了这样一个结果：如果两个紧致的类空超曲面Σ_1和Σ_2能通过一个紧致的时空M相衔接，并且在M上能定义明确的时间方向，且不包含闭合类时曲线，则Σ_1和Σ_2必定有相同的拓扑结构。霍金及同事也证明了一系列类似的结果。用通俗的话说，这些结果意味着在有界（“紧致”）的且具有良好因果性质（“能定义明确的时间方向，且不包含闭合类时曲线”）的时空中，拓扑改变是不可能的。除此之外，如果假定时空是全局双曲的，则还可以证明这样的结果：全局双曲时空必定可以分解为空间部分和时间部分，其中空间部分（即对应于任一时间坐标的类空超曲面）全都具有相同的拓扑结构。这同样意味着拓扑改变是不可能的（而且顺带保证了定义拓扑改变所需要的“将时空流形分解为空间部分和时间部分”是可以做到的）。

虽然这些结果所要求的条件不无苛刻性，且大都极难验证，但它们的存在无疑为构筑虫洞的可能性投下了巨大的阴影。这可以说是在我们前面介绍过的有关可穿越虫洞的诸多困难之外又增添了一种新困难，而且有可能是原则上不可克服的困难[36]。不过，这也并非是给虫洞物理学判处了死刑。因为即便这些结果成立，也还存在这样的可能性，那就是空间的初始结构已经包含虫洞。如果这样的话，我们（或“无限发达的文明”）所要做的只是把已经存在的虫洞“做大”，使之成为可穿越虫洞。而这——如我们在前面各节中所讨论的——虽然极端困难，却不是原则上不可能的。

关于虫洞的另一个讨论话题是，虫洞究竟是不是星际旅行的捷径？很多人——包括科普作家乃至物理学家——提到虫洞时都会想当然地认为那是星际旅行的捷径，比如像图5.3所示的那样。其实，虫洞并不一定是连接两个时空区域的捷径。图5.3本身就是对这一点很好的诠释，因为从图中可以看到，虫洞之所以成为捷径，完全是因为外部空间被弯曲成了U形，倘若没有这种堪称离奇的弯曲，虫洞就失去了最大的价值。事实上，即便在苹果这一原始比喻中，虫洞所能起到的捷径作用也

是十分有限的。倘若外部空间本身是接近平直的，则虫洞非但不一定是捷径，反而很可能是绕远。这一点几乎被有关虫洞的科幻、科普乃至专著所一致忽略，对于虫洞的应用来说却是一瓢不容忽视的冷水〔37〕。

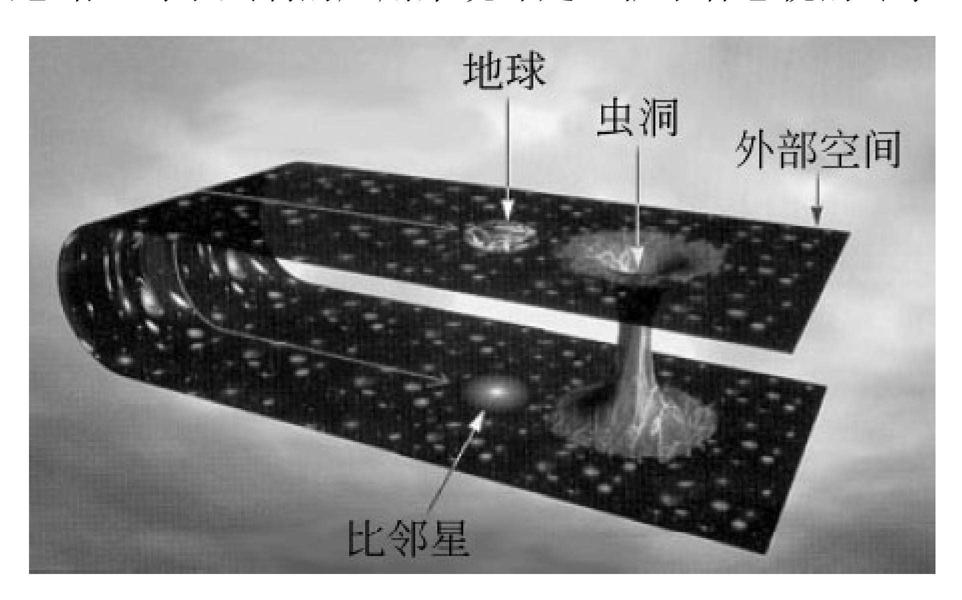


图**5.3**　典型的虫洞示意图

以上两个讨论话题一个为虫洞的构筑投下了阴影，一个为虫洞的应用泼下了冷水，对虫洞爱好者来说大概都不是喜闻乐见的。接下来讨论一个比较“正面”的话题吧。我们在上节的末尾曾经留下一个问题，那就是有关虫洞“工程学”的讨论及所得到的（很不乐观的）基本结论在更一般的情形下是否有可能绕过？对这个问题物理学家们也已做过一些探讨，得到的结果是勉强可算正面的。

比方说，在薄层近似下（这几乎是唯一能让我们摆脱球对称的近似），有物理学家研究了形状为立方体（或长方体）的虫洞，结果发现，这种虫洞的奇异物质只需分布在立方体（或长方体）的12条棱上即可。这意味着星际飞船可以不必与奇异物质直接接触就穿越虫洞，从而可以绕过最严峻的困难之一，即虫洞物质的张力有可能对星际飞船及乘员造成危害。有意思的是，比这一理论早得多，甚至早在虫洞这一概念因萨根与索恩的交流而在科幻小说中正式登场之前的1968年，著名英国

科幻作家克拉克（Arthur Clarke，1917—2008）就在名著《2001：太空奥德赛》（2001：A Space Odyssey）中引进了一个他称之为“星之门”（stargate）的概念。这个概念与虫洞的出入口极为相似。不仅如此，克拉克引进的“星之门”的形状恰好也是长方体，与几十年后的虫洞研究形成了有趣的呼应〔38〕。

不过，形状为立方体（或长方体）的虫洞从概念上讲虽不存在困难，从“工程学”角度看却绝不容易，因为计算表明，它12条棱上的奇异物质的张力必须达到10^{43}牛顿（N）的惊人数值，约相当于普通原子组成的一维结构所能提供的张力大小的一万亿亿亿亿亿（10^{52}）倍。因此，这方面的结果虽属正面，却极为勉强，以可行性而论，与不那么正面的结果相比，恐怕只是五十步与一百步之别。

星际飞船及乘员将会遇到的另一种应力——潮汐力——的情况又如何呢？研究表明，那可以通过选择特殊结构的虫洞及限制飞船的速度而加以控制甚至部分消除，从而也是可以绕过的。这一点甚至无须计算就可看出。比如由式（5.5.7）给出的径向潮汐加速度可以通过选择φ为常数的虫洞而彻底消除。此时的横向潮汐加速度则因为正比于v^2，而可以通过限制飞船的速度加以控制。因此，潮汐力带来的困难是可以绕过的。但是，绕过张力困难（如果可能的话）与绕过潮汐力困难所需满足的条件是不同的，能否同时绕过两者则是一个未知数。

除设法绕过张力困难与潮汐力困难外，关于虫洞还有一种有趣的设想值得一提，那就是将它与所谓的生命传输机（transporter）结合起来。生命传输机是一种将人体分解为基本粒子进行传输，并在目的地予以复现的装置。这种装置在著名美国科幻电视连续剧《星际迷航》（Star Trek）中有着极频繁的使用，堪称“招牌”技术。由于组成生命的基本粒子本身并无独特性，生命传输机甚至可以简化为只传输复现人体所需的信息。倘若把这种技术与虫洞结合起来，那么通过虫洞传输的将只是组成人体或传递信息的基本粒子。这虽然也绝不是一件容易的事

情，但虫洞“工程学”上最令人头疼的“穿越过程中遇到的应力是人体能够承受的”这一要求，以及由此导致的有可能极为巨大的奇异物质数量，就也许不再是严重问题了，从而将大大增加利用虫洞进行星际旅行的可能性。不过，这种设想也有自己的问题，其中最大的问题是：生命传输机本身的可行性就是一个未知数，而且是很不容乐观的未知数〔39〕。

在结束本节——从而也结束本专题及全书——之前〔40〕，还有一条隐伏在前文的线索需要交待一下。我们在5.4节的末尾曾经提到过，奇异物质的存在离不开以特殊方式配置起来的环境条件，我们并且还提到了一种可能性，即那些由普通物质构成的环境条件有可能对奇异物质的“效用”产生重大干扰，甚至彻底地抵消和淹没之。从某种意义上讲，这种可能性已得到了一些新近研究的证实。

20世纪90年代，美国塔夫斯大学（Tufts University）的物理学家福特（Larry Ford）等人对奇异物质进行了研究，结果发现它的负能量密度会无可避免地受到正能量密度的抵消。并且他们证明了一个定量的结论，即负能量密度必须满足一个被称为“量子不等式”（quantum inequality）的约束条件。用数学公式来表示，所谓量子不等式是这样一个不等式：

$$\int \rho f(\tau)\,d\tau \geqslant -\frac{C}{(\Delta\tau)^4} \qquad (5.7.1)$$

其中ρ是负能量密度；f（τ）是一个对时间自变量τ的积分为1，且特征宽度为Δτ的“抽样函数”（sampling function）；C是一个常数（通常小于或远小于1）。这个不等式的物理意义是：负能量密度的数值越大，所能维持的时间就越短。

量子不等式有很大的普遍性，不仅适用于平直空间，也适用于弯曲空间，它的发现引起了不少物理学家的兴趣。这是因为我们在本书的前

面各专题——即奇点与奇点定理、正质量定理、宇宙监督假设——中介绍过的各种经典广义相对论的结果都依赖于能量条件，从而都会受到负能量密度的威胁。而量子不等式的存在似乎为这种威胁设下了限制。这种限制能否在一定程度上挽救那些经典结果？这是不少物理学家感兴趣的问题。不过，与负能量密度对那些其他经典结果构成威胁相反，虫洞物理学却偏偏是需要负能量密度的，因此为负能量密度设下限制的量子不等式对其他经典结果或许是福音，对虫洞物理学却很可能是凶兆。

这种凶兆究竟“凶”到什么程度呢？福特等人针对球对称虫洞的情形进行了估算，结果表明，在一般情况下，虫洞只能是微观的，其半径将会有一个约为10^{-31}米（相当于原子线度的十万亿亿分之一）的上限。那样的虫洞当然不仅绝不是任何星际飞船能够穿越的，就连与生命传输机相结合都会是很成问题的。不过，福特等人也并未将所有的希望都掐灭，在他们分析过的情形中，原则上也有结构非常特殊的大型虫洞可以存在，但要求是：奇异物质必须分布在厚度比原子核还小——甚至小得多——的薄层内。这样的分布在理论上是不陌生的（几乎就是薄层近似），但实际上如何实现却是彻底的未知数。其他研究者也做了种种估算，得到的结果总体上也是不乐观的，甚至互有争议。唯一可以确定的是，迄今尚不存在一个哪怕只是理论上有希望的虫洞解，能解决前面提到过的所有困难。因此，在虫洞这个课题上，幻想与现实之间的距离还是相当遥远的。

本书到这里就完全结束了。最后要提醒读者的是，本书所介绍的各个专题都还在继续发展之中。很多细节可能还有推敲余地，很多结果可能还有改进空间，很多近似还有待突破，很多线索还有待探讨。所有这些都是让物理学家们每天清晨满怀着兴趣和希望走进办公室的动力之一，也是值得年轻人思考和参与的课题。广义相对论是一个优美的理论，优美得足以让人产生再无研究可做的错觉，如果本书能使某些读者注意到这个优美理论的某些尚有可为的前沿课题，甚至产生兴趣，那作

者的努力就得到了最好的回报。

注释

〔1〕惠勒不仅是索恩的导师，还曾与索恩及美国物理学家米斯纳（惠勒的另一位学生）合作撰写了有广义相对论“圣经”之称的著名教材《引力》（Gravitation）。

〔2〕因为这一原因，这一由施瓦西度规变形而来的爱因斯坦-罗森桥有时也被冠以虫洞之名，称为施瓦西虫洞（Schwarzschild wormhole）。

〔3〕这一点常常被过分强调，但其实潮汐力的大小是随黑洞质量的增加而减小的，因此只要黑洞质量足够大，潮汐力是可以变得足够小的。比如对于施瓦西黑洞来说，只要质量在几万个太阳质量以上，视界附近的潮汐力就能减小到人体所能承受的程度。这种质量虽绝非单个恒星的引力坍塌所能达到，对于据信普遍存在于星系中心的巨型黑洞来说却是小菜一碟。因此，潮汐力作为一种困难，并没有通常渲染的那样严峻。当然，许多科幻小说中的星际旅行通道的出入口往往建在行星上，或行星附近，那样的设想与上述巨型黑洞倒确实是完全不相容的。

〔4〕当然，科幻作品中还有其他一些有关星际旅行的设想，比如科幻作家阿西莫夫（Isaac Asimov，1920—1992）在著名的银河帝国系列（Galactic Empire series）和基地系列（Foundation series）等作品中就频繁采用了所谓“跃迁”的方式来进行星际旅行。不过这种“跃迁”背后的理论基础近于空白。另外还有一种很流行的设想是所谓的曲速引擎（warp drive），限于篇幅我们不作介绍了。

〔5〕“普朗克长度”是由万有引力常数（gravitational constant）、普朗克常数（Planck's constant）及光速（speed of light）构成的具有长度量纲的量，数量级为10^{-35}m。

〔6〕这里还有一点可以补充（或作为这一理解的应用），那就是可穿越虫洞的出入口所连接的渐近平直时空既可以位于同一个宇宙中，也可以位于不同宇宙中，前者被称为“宇宙内”（intra-universe）可穿越虫洞，后者被称为“宇宙间”（inter-universe）可穿越虫洞。这两种可穿越虫洞的主要区别在于时空的大尺度拓扑结构，而虫洞本身的结构可以视为相同，因此在我们的讨论中将不予区分。

〔7〕我们对条件的罗列与索恩和莫里斯的原始论文有所不同，但实质是一致的。

〔8〕依据这一物理意义，$b_+(\infty)$和$b_-(\infty)$的有限性可视为条件（5）（物质的数量是

可观测宇宙可以提供的）的要求。

〔9〕粗看起来，$e^{-2\varphi}$（1－b/r）在“喉咙”处为零是式（5.3.2）的推论。不过，它其实还有赖于前面提到的另一个结果，请读者想一想是什么结果。

〔10〕忘记了什么是“全局双曲时空”的读者请参阅2.4节。

〔11〕“奇异物质”这一术语有时还被用来表示其他含义，比如由所谓“奇异重子”（exotic baryon）组成的物质等。在本书中，我们只用它来表示破坏零能量条件的物质。

〔12〕感兴趣的读者可参阅，比如，依捷克森（Claude Itzykson）和祖柏尔（Jean-Bernard Zuber）的《量子场论》（Quantum Field Theory）。该书有中文版，科学出版社1986年出版。

〔13〕对物理学中的近似有一定了解的读者都知道，物理学中所谓的“大”和“小”乃是相对于具体问题的需要而言的，因此“平行导体板足够大”并不意味着它们看上去一定是庞然大物。在卡西米尔效应的研究中，平行导体板的“大”乃是相对于间距而言的，假如间距非常小（比如只有几微米甚至更小），则看上去很小（比如面积只有几平方厘米）的导体板也可以被认为是“足够大”的。

〔14〕确切地说，ħ是所谓“约化普朗克常数”（reduced Planck's constant），它与德国物理学家普朗克（Max Planck，1858—1947）最早引进的“正牌”普朗克常数h的关系是ħ＝h / 2π。不过约化普朗克常数在量子理论中的运用早已达到了“喧宾夺主”的程度，以至于直接称其为普朗克常数也未尝不可了。

〔15〕这一结果无须复杂的额外计算就可得到——只要考虑到洛伦兹协变性要求T^{ab}只依赖于η^{ab}和$z^a z^b$（z为导体板的法向单位矢量），以及电磁场的能量动量张量满足$T^{ab}\eta_{ab}=0$这一特殊性质即可。

〔16〕由于这一工作，玻德有时也被视为卡西米尔效应的共同发现者。

〔17〕这类验证中最早的一个是华盛顿大学（University of Washington）的物理学家拉莫理奥克斯（Steve Lamoreaux）于1997年发表的（实验完成于1996年），针对的是一块导体板与一个导体球之间的卡西米尔效应（从而避免了让两块导体板靠得极近而又互不接触的困难），验证精度约为5％。

〔18〕对于平行导体板情形，事实上可以很容易地证明，构成环境条件的普通物质（即导体板）的总能量远远超过依赖于它们而存在的负能量的数量。感兴趣的读者不妨利用本节给出

的有关结果，以及关于普通物质的某些“常识”，自行证明一下。

〔19〕这里我们虽然强调了“喉咙附近”，但球对称可穿越虫洞的喉咙所在处本身零能量条件也很可能是遭到破坏的，只不过是不能完全排除刚好处于破坏边缘这一特殊情形而已。但这种理论上的边缘情形在实际上出现的概率是可以忽略的。

〔20〕物质在原子线度上所能承受的张力可以用很简单的方法来估算，那就是物质在原子线度上的特征能量除以原子的体积，前者约为10^{-19}J（即电子伏特量级），后者约为$10^{-30}m^3$（即线度约为0.1nm），相应的张力约为$10^{11}J/m^3$，与$5\times10^{10}J/m^3$相当接近。

〔21〕细心的读者也许注意到了，这里我们用了“如果”一词。这是因为虽然一般说来，当物体穿越一种具有内部张力的结构时，自身也会受到与那种张力同一量级的张力作用，但考虑到我们对奇异物质的性质还所知不多，不能完全排除它与普通物质的相互作用特别弱，从而使穿越它的普通物体所受张力远小于奇异物质内部张力之类的可能性，因此在措辞上特意留有余地。

〔22〕不过，虽然索恩将此类问题冠上了萨根的大名，萨根本人却似乎从未直接用过“无限发达的文明”这一概念。他倒是沿袭拓展过苏联天文学家卡达什耶夫（Nikolai Kardashev，1935— ）提出的将星际文明分为Ⅰ类、Ⅱ类、Ⅲ类的所谓“卡达什耶夫标度”（Kardashev scale）。

〔23〕ε“很小”的具体含义是$\varepsilon \ll M^2$（感兴趣的读者可以试着恢复这一关系式中的物理常数）。

〔24〕感兴趣的读者可以计算一下这一度规所对应的能量动量张量，证实零能量条件会遭到破坏，且喉咙处的张力满足式（5.3.5）。关于近施瓦西虫洞，一个有趣的问题是：由于它的度规与施瓦西度规很相似，而施瓦西度规所描述的天体被认为是普遍存在的，这是否意味着近施瓦西虫洞可以通过对普通天体做适当的“手术”而比较容易地构筑出来？可惜，这个问题尚无答案，因为近施瓦西虫洞的度规与施瓦西度规虽然相似，时空的整体性质却很不相同。在这种情形下，由一种时空变为另一种时空别说难易，就连是否有可能都还是一个未知数（关于后面这一点，后文会有进一步的讨论）。不过，近施瓦西虫洞作为一个容易计算的例子，无疑是很有用的。

〔25〕如果把虫洞喉咙附近奇异物质的张力可能起到的破坏作用也考虑进去的话，那么飞船及乘员的残骸还可能被进一步撕裂，由虫洞另一端飞出的可能将是一串无法分辨来源的亚原

子粒子！

〔26〕这里所谓的“在某种意义上”是指从外部时空的角度来看。

〔27〕读者们想必还记得，索恩的“第一次”检讨是在收到昔日学生佩奇的来信，提醒他可穿越虫洞破坏零能量条件这一结论能通过全局方法得到时所作的（参阅5.3节）。不过从两次检讨本身的时间顺序上讲，其实这次在先，那次在后。

〔28〕这里的L是虫洞的入口和出口之间沿外部空间（即不通过虫洞）所测得的距离。“等时连接”则是指进入虫洞和离开虫洞的时刻都可以视为t（这样做的条件是虫洞的长度——从而穿越虫洞所需的时间——相比于L可以忽略）。

〔29〕这种利用狭义相对论的时钟延缓效应来产生时间差的方式在历史上曾引起过争议，并被冠以“时钟佯谬”之名，但实际上并无佯谬可言。

〔30〕这一步并非必需，而且并非一定得让距离L接近于零。事实上，距离L只要小于时间差T就可以了（感兴趣的读者请想一想，这是为什么？另外也请想一想，距离L“缓慢地缩小”中“缓慢”二字的用意何在？）

〔31〕不过，从构造步骤中不难看出，这种时间机器所能让人重返的“过去”不是完全任意的，它必须晚于虫洞被构筑成时间机器的时刻。这一结论——即时间机器不能让人重返它被构筑之前的年代——被认为有可能是普遍成立的。这一结论不仅有很强的“现实”意义，而且还有很重要的理论意义，因为霍金曾经提出过一个针对时间机器的有趣的诘难，那就是：倘若时间机器的构筑是可能的，那为什么我们从未遇到过来自未来世界的时间旅行家呢？这个诘难的潜台词是：我们从未遇到过来自未来世界的时间旅行家，最有可能的原因是时间机器在整个时间长河中——也就是永远——都不曾被构筑过，而这背后最有可能的原因则是时间机器的构筑是不可能的。但假如时间机器不能让人重返它被构筑之前的年代，霍金的诘难就不攻自破了，我们也就不一定要得出这么悲观的猜测了。

〔32〕用虫洞构筑时间机器除了是“虫洞”与“时间机器”这两个热门概念的“强强组合”外，还有一个引人注目的优点，那就是霍金等人在对时间机器进行研究后曾经发现：如果能量密度处处非负，则在任何有限时空区域内建造时间机器的努力都会产生奇点。这一结论被认为是对时间机器的一种严峻挑战。但用虫洞构筑的时间机器却是一个例外，因为虫洞的引进破坏了“能量密度处处非负”这一前提条件。

〔33〕这种悖论中最典型的是所谓的“祖父悖论”（grandfather paradox），即时间旅行家重返过去后设法阻止自己的祖父母相识。这个悖论的“悖”理在于：时间旅行家的祖父母若无法相识，他的父亲就不会出生，从而也就不会有时间旅行家自己了。但时间旅行家自己若不存在，又怎能重返过去并阻止祖父母相识呢？

〔34〕对于这种可能性，索恩本人持有一定的异议，他认为虫洞喉咙处的奇异物质会因负能量的缘故，而对辐射产生散焦（defocus）作用。这种散焦作用可以显著削弱辐射与它自身相叠加的效果，从而避免时间机器被摧毁。这方面——以及更一般的有关“时序保护假设”——的研究目前仍在进行之中。

〔35〕细心的读者也许注意到了，我们此处遇到的问题及解决问题的方法与3.2节所遇到的定义广义相对论动力学的问题及解决的方法是完全相似的。

〔36〕读者可能会问：这些理论结果都是经典广义相对论中的结果，量子引力有可能会推翻这些结果吗？这一问题的答案物理学家们也很想知道，可惜迄今为止还没有人能够确知（因为令人满意的量子引力理论本身尚不存在）。不过有迹象表明量子引力理论有可能不会推翻这些结果，比如在量子引力理论候选方案之一的正则量子化方案中，位形空间是由经典广义相对论确定的，从而会自动遵循后者有关拓扑改变的理论结果。

〔37〕当然，将外部空间弯曲成U形也许不是原则上不可能办到的事情，从而对于“无限发达的文明”来说也许不算问题。

〔38〕不过，克拉克的“星之门”并无理论基础，以长方体作为形状也“纯属巧合”，估计仅仅是为了让它看起来更像“人工”之物。

〔39〕关于这方面的介绍，可参阅拙作“生命传输机”（《科学画报》2003年第10期，上海科学技术出版社）。

〔40〕细心的读者也许会指出：5.3节列出的可穿越虫洞所需满足的条件中，“在微扰下保持稳定”这一条（条件（6））尚未介绍过。对此我们补述一下。这一条的直接研究是非常困难的，因为无可避免地要面对非球对称的情形。不过，索恩和莫里斯的原始论文对这一条做了一个很有道理的评注，那就是即便虫洞是不稳定的，一个“无限发达的文明”仍可以通过密切监视它的结构，而将不稳定的苗头随时扼杀在摇篮里。因此，可穿越虫洞的“维稳”不存在原则上的困难。

名词索引

A

B

C

超引力
潮汐力
虫洞
虫洞物理学
抽样函数

D

大爆炸
狄拉克方程
狄拉克矩阵

F

范德瓦尔斯力
非类空测地线
非时序点集
菲尔茨奖
费米子
封闭陷获面
弗里德曼解
负能量物质

G

高斯-柯达西方程组
共轭点
共形变换
共形紧致化
共形无穷远
孤立体系

广义仿射参数
过去柯西视界
过去柯西展开

H

哈密顿表述
黑洞
黑洞唯一性定理
黑洞无毛发定理
霍金-彭罗斯奇点定理

J

奇点
奇点定理
迹能量条件
几何动力学
渐近平直时空
近施瓦西虫洞

K

卡西米尔力
卡西米尔效应
柯西面
柯西展开
壳层穿越奇点
壳层会聚奇点
可穿越虫洞
克尔-纽曼度规

克尔-纽曼黑洞

克尔-纽曼解

克莱茨曼标量

克莱因-戈登场

L

雷查德利方程

雷斯勒-诺斯特朗姆度规

雷斯勒-诺斯特朗姆黑洞

雷斯勒-诺斯特朗姆解

类光测地线

类光一般性条件

类空测地线

类空超曲面

类时测地线

类时一般性条件

黎曼-彭罗斯猜想

黎曼联络

李导数

理想尘埃

理想流体

量纲分析法

量子不等式

量子涨落

临界参数

零尘埃

零能量条件

罗伯逊-沃尔克度规
裸奇点
洛伦兹度规
洛伦兹因子

M

闵科夫斯基度规

N

能量动量张量
能量条件

P

彭罗斯不等式
彭罗斯猜想
膨胀标量
平均能量条件
平均外曲率
普朗克常数

Q

奇异物质
强渐近可预测
强能量条件
强因果性条件
切变张量
曲率张量
全局方法

无限发达的文明

Y

雅可比场
一般性条件
因果性条件
引力微子
诱导度规
宇宙监督假设
宇宙学奇点

Z

张力
正则形式
正质量定理
终极理论
逐点能量条件
主能量条件
祖父悖论
最大柯西展开

人名索引

A

B

D

F

G

勒梅特（Georges Lemaître）
雷波维茨（Clement Leibovitz）
雷查德利（Amal Raychaudhuri）
栗弗席兹（Evgeny Lifshitz）
罗伯逊（Howard Robertson）
罗曼（Thomas Roman）
罗森（Nathan Rosen）
洛卡姆（Joachim Lohkamp）

M

马斯登（Jerrold Marsden）
米斯纳（Charles Misner）
莫勒（Christian Møller）
莫里斯（Mike Morris）
莫里斯奇（Osvaldo Moreschi）

N

纳里卡（Jayant Narlikar）

P

帕克（Thomas Parker）
佩奇（Don Page）
彭罗斯（Roger Penrose）
普赖斯（Richard Price）
普雷斯基尔（John Preskill）

Q

钱德拉塞卡（Subrahmanyan Chandrasekhar）

乔希（Pankaj Joshi）
丘成桐（Shing-Tung Yau）

S

萨根（Carl Sagan）
塞弗特（Hans Seifert）
森（Ashoke Sen）
舍恩（Richard Schoen）
施奈德（Hartland Snyder）
施瓦西（Karl Schwarzschild）
斯帕林（George Sparling）
斯图尔特（John Stewart）
索恩（Kip Thorne）
索金（Rafael Sorkin）

T

泰勒（Joseph Taylor）
泰特伯姆（Claudio Teitelboim）
陶布斯（Clifford Taubes）
梯普勒（Frank Tipler）
托曼（Richard Tolman）

W

威顿（Edward Witten）
威尔斯（H. G. Wells）
维希外希瓦拉（C. V. Vishveshwara）
沃尔德（Robert Wald）
沃尔克（Arthur Walker）

伍尔加（Eric Woolgar）

X

肖盖（Yvonne Choquet-Bruhat）
肖普推克（Matthew Choptuik）

Y

伊尔玛能（Tom Ilmanen）
伊斯雷尔（Werner Israel）
尤瑟福（Ulvi Yurtsever）
约德杰斯（Peter Yodzis）
约尔当（Pascual Jordan）

参考文献

综合读物

1. Ciufolini I, Wheeler J A, Gravitation and Inertia [M]. Princeton: Princeton University Press, 1995.
2. Hawking S W, Ellis G F R. The Large Scale Structure of Space-Time [M]. Cambridge: Cambridge University Press, 1975.
3. Joshi P S. Global Aspects in Gravitation and Cosmology [M]. Oxford: Oxford University Press Inc., 1993.
4. Misner C W, Thorne K S, Wheeler J A. Gravitation [M]. New York: W. H. Freeman and Company, 1973.
5. Wald R M. General Relativity [M]. Chicago: University of Chicago Press, 1984.
6. Witten L. Gravitation: An Introduction to Current Research [M]. New York: Wiley 1962.
7. 刘辽，赵峥．广义相对论[M]．北京：高等教育出版社，2004.

能量条件

1. Barcelo C, Visser M. Twilight for the energy conditions? [J]. Int. J. Mod. Phys, 2002, D11: 1553-1560.
2. Ford L H. The Classical Singularity Theorems and their Quantum Loopholes [J]. Int. J. Theor. Phys., 2003, 42: 1219-1227.

奇点与奇点定理

1. Foertsch T, Wegner B. Selected Topics in Singularities of Space-time, Proc. Summer School on Diff [C]. Geometry-Dep. de Matematica, Universidade de Coimbra-September 1999: 39-52.
2. Hawking S W, Penrose R. The Nature of Space and Time [M]. Princeton:

Princeton University Press, 1996.
3. Natário J. Relativity and Singularities-A Short Introduction for Mathematicians [J]. arXiv: math/0603190[math.DG].
4. Schoen R, Yau S T. The Existence of a Black Hole Due to Condensation of Matter [J]. Comm. Math. Phys., 1983, 90: 575-579.
5. Senovilla J M. Singularity Theorems in General Relativity: Achievements and Open Questions [J]. arXiv: physics/0605007[physics.hist-ph].
正质量定理
1. Bray H L. Proof of the Riemannian Penrose Inequality Using the Positive Mass Theorem [J]. J. Diff. Geom., 2001, 59: 177-267.
2. Bray H L, Lee D A. On the Riemannian Penrose Inequality in Dimensions Less Than 8 [J]. arXiv: 0705.1128 [math.DG］.
3. Frauendiener J. Conformal Infinity [J]. Living Rev. Relativity, 2004, 7: 1.
4. Gibbons G W, et al. Positive Mass Theorems for Black Holes [J]. Comm. Math. Phys., 1983, 88: 295-308.
5. Kazdan J L. Positive Energy in General Relativity [J]. Séminaire N. Bourbaki, 1982-1983，25: 315-330.
6. Lohkamp J. The Higher Dimensional Positive Mass Theorem Ⅰ[J]. math.DG/0608795.
7. Moreschi M, Sparling G A J. On the Positive Energy Theorem Involving Mass and Electromagnetic Charges [J]. Comm. Math. Phys., 1984, 95: 113-120.
8. Parker T, Taubes C H. On Witten's Proof of the Positive Energy Theorem [J]. Comm. Math. Phys., 1982, 84: 223-238.
9. Penrose R, Sorkin R D, Woolgar E. A Positive Mass Theorem Based on the Focusing and Retardation of Null Geodesics [J]. gr-qc/9301015.
10. Schoen R, Yau S T. On the Proof of the Positive Mass Conjecture in

General Relativity [J]. Comm. Math. Phys., 1979, 65: 45-76.
11. Witten E. A New Proof of the Positive Energy Theorem [J]. Comm. Math. Phys., 1981, 80: 381-402.
12．丘成桐，孙理察．微分几何讲义[M]．北京：高等教育出版社，2004.
宇宙监督假设
1. Choptuik M W. Universality and Scaling in Gravitational Collapse of a Massless Scalar Field [J]. Phys. Rev. Lett., 1993, 70: 9.
2. Harada T, Iguchi H, Nakao K. Physical Processes in Naked Singularity Formation [J]. Prog. Theor. Phys., 2002: 107: 449.
3. Joshi P S. Cosmic Censorship: A Current Perspective [J]. Mod. Phys. Letts., 2002, A17: 1067.
4. Mars M. An Overview on the Penrose Inequality [J]. Journal of Physics: Conference Series, 2007, 20: 012004.
5. Penrose R. The Question of Cosmic Censorship [J]. J. Astrophys. Astr., 1999, 20: 233-248.
6. Schoen R, Yau S T. The Existence of a Black Hole due to Condensation of Matter [J]. Comm. Math. Phys., 1983, 90: 575-579.
7. Singh T P. Gravitational Collapse, Black Holes and Naked Singularities [J]. arXiv: gr-qc/98050661.
8. Wald R M. Gedanken Experiments to Destroy a Black Hole [J]. Ann. Phys., 1974, 83: 548-556.
9. Wald R M. Gravitational Collapse and Cosmic Censorship [J]. arXiv: gr-qc/9710068.
虫洞物理学
1. Bordag M (eds). The Casimir Effect 50 Years later [M]. Singapore: World Scientific, 1999.

2. Dalvit D, Milonni P, Roberts D, Rosa F (eds). Casimir Physics [M]. Berlin Heidelberg: Springer, 2011.
3. Ford L H, Roman T A. Quantum Field Theory Constrains Traversable Wormhole Geometries [J]. Phys. Rev., 1996, D53: 5496-5507.
4. Halpern P. Cosmic Wormholes: The Search for Interstellar Shortcuts [M]. New York: Plume, 1993.
5. Hawking S W. Hawking on the Big Bang and Black Holes [M]. Singapore: World Scientific, 1993.
6. Lamoreaux S K. Demonstration of the Casimir Force in the 0.6 to 6 μm Range [J]. Phys. Rev. Lett. 1997, 78: 5.
7. Milton K A. The Casimir Effect: Physical Manifestations of Zero-Point Energy [M]. Singapore: World Scientific, 2001.
8. Morris M S, Thorne K S. Wormholes in spacetime and their use for interstellar travel: A tool for teaching general relativity [J]. Am. J. Phys. 1988: 56: 395.
9. Roman T A. Some Thoughts on Energy Conditions and Wormholes [J]. arXiv: gr-qc/0409090.
10. Thorne K S. Black Holes & Time Warps [M]. New York: W. W. Norton & Company, 1995.
11. Visser M. Lorentzian Wormholes: From Einstein to Hawking [M]. Woodbury: AIP Press, 1996.

后　　记

终于到了可以写后记的时候，这大概是写一本书的过程中最轻松的时刻吧。这本书前后拖了七年多，跟上一本书——《黎曼猜想漫谈》——相比虽还算“高效”，也终究够得上写点回忆了——毕竟，人生能有几个“七年多”呢？

本书的缘起可以追溯到我写科普的起点：2002年。那一年，我开始撰写一个题为星际旅行漫谈的科普系列。那个系列承袭了我在天文和科幻两方面的兴趣，是我最早的科普。其中有关虫洞的部分则可以说是本书的先声。记得很早的时候——也许是还在念中学的那会儿吧——我曾从一本科普书上读到过一位科学家（或科普作家）的话，大意是：光速是有限的，我们的子孙后代将在极度的孤寂中体会它的深意。那句话给我的印象非常深（可惜却记不起出处了，只能把惭愧推给年岁），在撰写星际旅行漫谈时，也常常会想起它来。我喜爱物理，但物理规律有时连我自己也觉得有些冷——或者所有真正的规律都是冷峻的吧，想到浩瀚星空中被光速极限无情分隔开的遥远部分，是我们几乎永远也不可能有机会前往的，不免有些惆怅。

在那样的背景下，虫洞这种似乎能变相突破光速极限的手段引起了我的极大兴趣。而且跟其他一些科幻设想——比如阿西莫夫笔下的“跃迁”（jump），或某些天知道怎么造出来的“超光速飞船”——相比，虫洞的物理色彩要浓厚得多，甚至可以算是有一定理论根基的，这对于我这种既喜欢科幻，又是物理专业的人来说，是格外有吸引力的。当然，如果您已经读过本书，那么也许已经跟我一样不无遗憾地注意到了，依据我们现在所知的物理规律，虫洞作为星际旅行的通道，实现的可能性是微乎其微的。作为科幻爱好者，我期盼正面消息，甚至梦想着在自己有生之年就能听到那样的消息；但作为科普作者，我只能忠实于现代物

理，那是确立科普作品价值的核心坐标系，也是它有别于幻想的存在理由。当然，凡存在显著争议的部分，我都尽量给出了提示——不仅为了严谨，也是因为那是某种程度的希望所系，是我自己所乐意见到的。

在虫洞这个题材上，我虽然已写过文章，但前面提到的星际旅行漫谈中有关虫洞的部分，以及后来发表在《科幻世界》上的“虫洞：旅行家的天堂还是探险者的地狱？”（《科幻世界》2006年第5期，科幻世界杂志社）都是普通科普。而要想把虫洞介绍得透彻些，显示出它在现代物理中的理论根基，则必须采用所谓专业科普的形式，这使我萌生了撰写本书的念头——也就是本书的缘起。

不过，撰写本书的念头虽来自对虫洞的兴趣，本书的题材却并不仅限于虫洞。因为在收集素材的过程中，我注意到在虫洞物理学中有一组条件起着很重要的作用，那就是能量条件，而能量条件在其他一些广义相对论课题——比如奇点与奇点定理、正质量定理、宇宙监督假设——中也有重要应用，那些题材也很有意思，并且大都是普通广义相对论教材中不介绍或介绍得不深入，从而对专业科普来说可算是好课题的。于是我决定扩展写作范围，以能量条件为线索，对那些课题一并进行介绍。由于这一原因，本书最初在我的网站上连载时，所用的标题乃是“现代物理中的能量条件”。同时，也是由于这一原因，本书的篇幅才膨胀为了“书”。

本书的撰写始于2005年12月，前两章的内容（即“能量条件”和“奇点与奇点定理”）曾于2007—2008年间在《现代物理知识》上连载。可惜，该连载只进行了两章，第3章（即“正质量定理”）因“不够通俗”而遭审稿编辑“枪毙”，整个连载遂告“腰斩”。当时我网站上的连载已进行到了第4章（即“宇宙监督假设”），由于约稿压力的终止，在2008年9月写完该章之后，我就停了下来，将写作兴趣转向了其他方面。于是本书也像《黎曼猜想漫谈》一样，一度成为了“烂尾楼”。

本书的重新启动是在2012年4月，当时《黎曼猜想漫谈》已经写

完，在与清华大学出版社的编辑讨论下一本书的计划时，我提到了这个已经完成了大半的题材。编辑表示感兴趣，并寄来了出版合同。于是自2012年12月起，我开始撰写本书的最后一章（即“虫洞物理学”），并于2013年3月完成了全书。

本书作为我的第四本书，已不再有出版第一本书那样的新鲜感，但它的完成仍有一个比较特殊的意义，那就是填补了一个不应有的空白。有些读者也许知道，我这位写过两本天文书、一本数学书的“非著名科普作家”其实是物理专业的。一个物理专业的人，写了三本书却没有一本是关于物理的，多少有些愧对专业，也有些缺憾。本书的完成可以算是弥补了缺憾，从这点上讲，它的出版对我有一种不同于前三本书的欣慰感。

当然，真正的欣慰不是我写了一本物理书，而应该是这本物理书能得到部分读者的喜爱——这一点我知道是不容易的，起码要比前三本书都更难，因为与同为专业科普的《黎曼猜想漫谈》不同，本书没有多少科学家的逸事，也没有什么戏剧性的情节，这对于博取读者的喜爱是非常不利的。但我希望，本书凭借“事后诸葛”特有的时间优势，能做到在叙述逻辑上比原始论文更顺畅或更易被理解，以及对后续发展有更清晰的视野。若如此，则它对那几个广义相对论专题可以起到导读作用。我并且希望，起码有一部分读者能从这样的导读中受益。

在结束这篇后记之前，有一点关于笔误的小小感想要在这里提一下。众所周知，在所有错误中，笔误是相对轻微的，往往读者自己就能分辨。但专业科普由于带有数学公式，笔误一旦出现在公式中，影响就会变大，且看上去不像笔误倒像硬伤。日本物理学家汤川秀树写过这样一句体会：“拿到出版的书后，每次都会有很多问题闯入我的眼帘，我会为此而坐卧不安。”我曾经觉得自己的脸皮要比汤川先生厚，不至于坐卧不安的，但后来在“豆瓣读书”看到一位读者针对《黎曼猜想漫谈》中一处公式笔误的留言“才看几页，就发现了公式错误，原本很期待的

一本书，很怀疑严谨性了”时感到脸上发热，才不无遗憾地意识到自己的脸皮其实也不厚。而更遗憾的是，那公式在我的网站上其实是正确的，纯属整理文稿时的疏漏——那种感觉，宛如昔日大考之后发现一个粗心大意的错误。为了减少笔误的影响，我在这里公布一下自己的网址，以便读者发现笔误或疑似笔误时，可以核对网络版，也可以发邮件告诉我。任何笔误一经确认，都将第一时间在网络版上更正，并将等待重印或再版的机会在书面版上更正。

最后，谢谢诸位的阅读，并请大家相信一位有着几十年读书史和十几年爬网史的“书虫”兼“网虫”的经验：阅读实体书的感觉绝对要比阅读网络版好得多。

作者网址：http://www.changhai.org/

2013年3月30日写于纽约

理解科学丛书·卢昌海科普著作



原点阅读

清华大学出版社

原点阅读（The Origin），清华大学出版社旗下的图书品牌，秉承“科学，让个人更智慧，让社会更理性”的理念，致力于科学普及和科技文化类图书的出版，传播科学知识、科学精神、科学方法，展现科学的真实、独立、智慧、多变、宽容、动人及迷人。